REACTIVITY AT SURFACES WITH HIGH-DIMENSIONAL NEURAL NETWORK POTENTIALS

Dissertation
for the award of the degree
“Doctor rerum naturalium” (Dr. rer. nat.)
of the Georg-August-Universität Göttingen
within the doctoral program Chemistry
of the Georg-August University School of Science (GAUSS)

submitted by
MARTIN LIEBETRAU
from Witzenhausen, Germany

Göttingen, 2023

Thesis Committee

Prof. Dr. Jörg Behler, Theoretische Chemie, Institut für Physikalische Chemie, Georg-August-Universität Göttingen

Prof. Dr. Alec M. Wodtke, Physikalische Chemie I, Institut für Physikalische Chemie, Georg-August-Universität Göttingen

Dr. Jörg Meyer, Leiden Institute of Chemistry, Leiden University

Members of the Examination Board

Reviewer: Prof. Dr. Jörg Behler, Theoretische Chemie, Institut für Physikalische Chemie, Georg-August-Universität Göttingen

Second Reviewer: Prof. Dr. Alec M. Wodtke, Physikalische Chemie I, Institut für Physikalische Chemie, Georg-August-Universität Göttingen

Further Members of the Examination Board:

Prof. Dr. Ricardo A. Mata, Computerchemie und Biochemie, Institut für Physikalische Chemie, Georg-August-Universität Göttingen

Prof. Dr. Burkhard Geil, Biophysikalische Chemie, Georg-August-Universität Göttingen

Jun.-Prof. Dr. Daniel Obenchain, Physikalische Chemie, Institut für Physikalische Chemie, Georg-August-Universität Göttingen

Jun.-Prof. Dr. Anna Krawczuk, Anorganische Chemie IV, Institut für Anorganische Chemie, Georg-August-Universität Göttingen

Date of the oral examination: June 2, 2023

Acknowledgments

First, I would like to thank Prof. Dr. Jörg Behler for being there to explain, to advice and to encourage me during this interesting project and the writing of this thesis. And especially for giving me the opportunity to do my master and doctor thesis in his group. Secondly i would like to thank Prof. Dr. Alec Wodtke subject and for being the second supervisor of my doctoral thesis. Additionally, I would like to thank the members of the examination board for taking the time for reading my thesis and for the oral exam. Many thanks go to all members of the Behler group for the pleasant and inspiring environment. I would like to thank my former bachelor and master students Maike Vahl and Lucas Kanagarajah and my colleagues and friends Marco, Dilshana, Xaiza, Alex, Moritz, Marius and Daniel for great discussions, the pleasant working atmosphere and the tons of fun we had. For the excellent technical support I thank Dr. Rainer Oswald who managed to solve all issues concerning the computer clusters.

I would also like to thank my good friend Mareike for the great time and vacations we spend together.

In addition, I would like to thank all members from the BENCh RTG 2455 for the support during my time in the training group and afterwards.

I would also like to thank my friends and my family for supporting me and keeping my mood up. Especially my girlfriend Alea for supporting me during the last few months!

Abstract

The surface of a material is critical for its properties, especially in heterogeneous catalysis. Surfaces properties can be probed, for example, by H atom scattering. However ab-inito theoretical investigations of H atom scattering are held back by the high cost for performing the large amount of necessary simulations. Machine learning potentials can be used to bridge the gap between the high accuray of ab-inito calculations and the low computational cost of force field calculations. In this thesis, a high-dimensional neural network potential (HDNNP) based on RPBE density funcitonal theory reference data has been constructed for the scattering of H atoms from the [0001] α-Al_2O_3 surface. This system is well accessible for theoretical studies and experimental benchmarks due to the simple nature of H atom scattering, lacking the involvement of steric or vibrational effects. The process of generating the reference data set and validating the constructed HDNNP, using system properties like the phonon band structure and the potential energy surface as a function of collective variables, is described in detail. Furthermore, the HDNNP is validated using experimental kinetic energy distributions and angular distributions from H atom scattering. Insights into the surface structure are gained by comparing the experimental and theoretical scattering distributions. Finally, the accuracy of the RPBE functional describing the interaction of the H atom with the surface is evaluated and discussed.

Contents

Acknowledgments **V**

Abstract **VII**

Abbreviations **XII**

1 Introduction **1**

2 Methods **4**

2.1 Density Functional Theory 4

2.1.1 Foundations of Density Functional Theory 4

2.1.2 Density Functional Theory 6

2.1.3 Spin-Polarised Calculations 11

2.1.4 Exchange Correlation Functionals 12

2.1.5 Periodic Boundary Conditions 13

2.1.6 Dispersion Correction 14

2.2 Machine Learning Potentials 15

2.2.1 Introduction to Machine Learning Potentials 15

2.2.2 High Dimensional Neural Network Potentials 16

2.2.3 Molecular Dynamics 22

2.2.4 Experimental Setup 23

3 Computational Details **24**

3.1 Density Functional Theory 24

3.1.1 Fritz Haber Institute-Ab Initio Materials Simulations 24

3.2 Molecular Dynamics Simulations using LAMMPS 24

3.2.1 Phonopy Calculations 25

3.3 High-Dimensional Neural Network Potentials 25

3.3.1 Construction of the Reference Data Set 25

3.3.2 RuNNer 38

4 Results and Discussion **39**

4.1 Validation of the Final RPBE α-Al_2O_3 High Dimensional Neural Network Potential 39

4.1.1 Overview 39

4.1.2 Reference Data Set 40
4.1.3 Validation of the bulk System using Density Functional Theory 41
4.1.4 Investigation of the H Atom Spin 43
4.1.5 Comparing PBE and RPBE H Atom Interactions 52
4.1.6 Validation of the Slab System using Density Functional Theory 55
4.2 Benchmarking the HDNNP with Experimental Data 63
4.2.1 Overview 63
4.2.2 Creation of Theoretical Kinetic Energy Distributions 64
4.2.3 Benchmarking Experimental and Theoretical Kinetic Energy Distributions 66
4.2.4 Benchmarking Angular Distributions 74
4.2.5 Discussion of the Differences between the Experiment and Theory 77

5 Conclusion **87**

6 Outlook **89**

References **90**

Appendix **97**
A.1 RuNNer Settings 97

Abbreviations

ACSF	Atom-Centered Symmetry Functions
CCSD(T)	Coupled Cluster Singles and Doubles with perturbative Triples
DFT	Density Functional Theory
FHI-aims)	Fritz-Haber-Institute ab initio materials simulations
FPS	farthest point sampling
GGA	Generalized Gradient Approximation
HDNNP	High-Dimensional Neural Network Potential
LAMMPS	Large-scale Atomic/Molecular Massively Parallel Simulator
LDA	Local Density Approximation
MC	Monte Carlo
MD	Molecular Dynamics
MLP	Machine Learning Potential
NH	Nosé-Hoover
PBC	Periodic Boundary Conditions
PES	Potential energy surface
RMSE	Root Mean Squared Error
SCF	Self-Consistent Field
TS	Tkatchenko-Scheffler
UHV	Ultra High Vacuum
vdW	van der Waals

1 Introduction

Heterogeneous catalysis is tremendously important in today's economy and state of the world. Catalyzed reaction contribute to a huge percentage of the world's gross domestic product.[1] Processes like the Haber-Bosch process for artificial ammonia synthesis changed the world by being the source of half of the nitrogen atoms in our bodies.[2] The surface of heterogeneous catalysts is critical for their properties. Scientists have made significant efforts to improve prominent heterogeneous catalysis processes by gaining a deeper understanding of their underlying chemical and physical dynamics. These studies have led to initial discoveries and subsequent industrial applications of improved catalysts.[3]

Theoretical studies of established and new materials and processes can further improve the understanding of catalytic reactions and chemistry in general.[4] The theoretical investigation of catalytic surfaces is difficult and cost intensive due to the presence of diverse morphologies and preparation methods. The usually high pressures and temperatures used in preparation further increase the complexity of the systems dramatically. However the rapid advancements in computer power, the continuous development of new electronic structure algorithms, and simulation setups that accurately replicate reaction conditions help to bridge the gap between industrial applications and theoretical simulations. One recently developed way to circumvent the high cost of the calculations is the use of machine learning potentials (MLP).[5–7]

Machine learning potentials, fitted to ab-initio reference calculations, can reduce the cost of simulations for large, long, or complex calculations, while retaining the high level of accuracy from the reference methods. They have been developed for many different systems, ranging from gas phase vibrational spectroscopy[8] to manganese oxide spinels used as electrode materials[9] to the description of solid-gas interfaces.[10]

Experimental benchmarks of these systems are very valuable to determine the accuracy of the models, but the increase in complexity of the theoretically described systems also increases the complexity of the needed experiments. Furthermore a machine learning potential is based on a data set consisting of tens of thousands of calculations which have to be consistent with each other. The choice of the reference method and its settings is of utmost important for a well working potential energy surface (PES). In order to keep the cost of generating such a data set low, the reference structures have to be selected carefully, so that a small data set still fully maps all relevant chemical environments. Thorough validation of the electronic structure reference calculations and the machine learning potential, as well as experimental results is needed for accurate and reliable results. The creation and validation of those models is not trivial and further research is still required. The investigation of less complex model systems can help understanding the complexity of creating machine learning potentials and also supports understanding what is needed from experimental benchmarks to ascertain the quality of the models.

Using H atom scattering to probe surface properties is a well-researched process. The scattering of H atoms and small molecules was studied from metal surfaces,[11] surface oxides,[12] graphene,[10] and insulators.[13] The interaction between an H atom and a well-defined surface in ultra-high vacuum (UHV) is a relatively simple process. Unlike other collisions, such as those involving vibrations or steric effects, only the atomic translational energy is involved, which makes it particularly attractive for research and comparison with first principle theories. Additionally, these interactions are important in a variety of technological applications, such as heterogeneous catalysis,[14] interstellar H_2 production on dust grains,[15] nuclear fusion in tokamaks, and hydrogen storage.[16,17]

While the experimental investigation is performed with highest precision, the theoretical studies are gatekept by the huge amount of required simulations. Machine learning potentials enable the simulation of millions of trajectories needed to investigate the surface and understand to the experimental results. One experimentally characterized system is the scattering of H atoms from α-Al_2O_3.[18] The [0001] α-Al_2O_3 surface is well suited for theoretical benchmarking. The surface termination is well studied under high vacuum conditions and does not show a complex surface reorganization. The material is also often applied, either as catalyst[19] or as supporting material[20] increasing the need of a deeper understanding of how this material interacts at interfaces.

In this thesis, a high-dimensional neural network potential[21] (HDNNP) is trained on RPBE density functional theory (DFT) data of the α-Al_2O_3 and H atom system. The structural and electronic setup is investigated and the accuracy of the RPBE functional is compared to the PBE functional. Using the bulk properties of α-Al_2O_3, including the phonon band structure, ab-initio DFT calculations including ab-inito molecular dynamics (MD) simulations and experimental measurements the HDNNP is thoroughly validated. MD trajectories for four different experimental incidient conditions are calculated using the HDNNP and the kinetic energy distributions for different scattering angles and angular distributions are compared to the experiment. The experimental surface termination is evaluated using the theoretical kinetic energy distributions at different incident azimuth angles. Differences between the experimental and theoretical surface structure and roughness are analyzed. The accuracy of the RPBE functional, describing the interaction of the H atom with the surface, are discussed and evaluated.

2 Methods

2.1 Density Functional Theory

2.1.1 Foundations of Density Functional Theory

According to the first postulate of quantum mechanics, the wave function $\Psi(\boldsymbol{\tau}, t)$ fully describes the state of a system, where the positions $\boldsymbol{r}$ and spins s of all particles are expressed by $\boldsymbol{\tau}$ at time t.[22] Although the wave function itself has no direct physical interpretation, the squared absolute value $|\Psi(\boldsymbol{\tau}, t)|^2$ corresponds to the probability density of the described system at a given time.[23] For a many-electron system, the wave function must be antisymmetric under the interchange of any two electrons, which is known as the Pauli principle:[24,25]

$$\Psi(\boldsymbol{\tau}_1, ..., \boldsymbol{\tau}_i, \boldsymbol{\tau}_j, ..., \boldsymbol{\tau}_n) = -\Psi(\boldsymbol{\tau}_1, ..., \boldsymbol{\tau}_j, \boldsymbol{\tau}_i, ..., \boldsymbol{\tau}_n). \tag{2.1}$$

A non-relativistic description of the system's time evolution can be obtained from the time-dependent Schrödinger equation:[26,27]

$$\hat{H}(\boldsymbol{\tau}, t)\Psi(\boldsymbol{\tau}, t) = i\hbar \frac{\partial \Psi(\boldsymbol{\tau}, t)}{\partial t} \ , \tag{2.2}$$

where $\hat{H}$ is the Hamiltonian operator and $\hbar$ is the reduced Planck constant. For a time-independent Hamiltonian, the space and time dependence of the wave function can be separated, yielding the time-independent Schrödinger equation:

$$\hat{H}(\boldsymbol{\tau})\Psi(\boldsymbol{\tau}) = E\Psi(\boldsymbol{\tau}) \ , \tag{2.3}$$

where E represents the energy of stationary eigenstates. The probability density of the particles is given by the square of the absolute value of the wave function $|\Psi(\boldsymbol{\tau})|^2$.[23] The Hamiltonian, defining the interaction of the electrons and nuclei, of N electrons and M nuclei is

$$\hat{H} = \underbrace{-\frac{1}{2}\sum_{i=1}^{N} \nabla_i^2}_{T_\mathrm{e}} \underbrace{-\sum_{A=1}^{M} \frac{1}{2M_A}\nabla_A^2}_{T_\mathrm{N}} \underbrace{-\sum_{i=1}^{N}\sum_{A=1}^{M} \frac{Z_A}{r_{iA}}}_{V_\mathrm{Ne}} \underbrace{+\sum_{i=1}^{N}\sum_{j>1}^{N} \frac{1}{r_{ij}}}_{V_\mathrm{ee}} \underbrace{+\sum_{A=1}^{M}\sum_{B>A}^{M} \frac{Z_A Z_B}{R_{AB}}}_{V_\mathrm{NN}} \ , \tag{2.4}$$

were r_{iA} represents the distance from electron i to nucleus A, r_{ij} denotes the distance between electron i and electron j, and R_{AB} signifies the distance between nucleus A and nucleus B. Nucleus A has a charge of Z_A and a mass of M_A. The Laplacian operators ∇_i and ∇_A correspond to electron i and nucleus A, respectively, and are defined in terms of the coordinates of the particles.

The Hamiltonian in equation 2.4 has several contributions. The first and second terms, T_e and T_N, represent the kinetic energies of the electrons and nuclei, respectively. The remaining terms, V_Ne, V_ee, and V_NN, describe Coulomb interactions between electrons and nuclei, with their signs indicating whether the interactions are attractive or repulsive. The third term, V_Ne, represents the Coulomb attraction between electrons and nuclei, while the fourth and fifth terms, V_ee and V_NN, express the repulsion between electrons and between nuclei.

The mass of an electron is significantly smaller than that of a nucleus, resulting in a much larger acceleration of electrons compared to nuclei. Since the motion of electrons and nuclei occurs on significantly different time scales, the two terms can be separated using the time independent Schrödinger equation. The electronic energy can be characterized using the electronic Hamiltonian H_elec, given by:

$$H_\mathrm{elec} = T_\mathrm{e} + T_\mathrm{Ne} + V_\mathrm{ee} \ . \tag{2.5}$$

This Hamiltonian in equation 2.5 describes the energy of electrons in the presence of fixed point charges corresponding to the nuclei. Since the nuclei are stationary, their kinetic energy is zero, and the energy of repulsion between the nuclei is constant. When a constant term is added to an operator, it only affects the operator eigenvalue and not the eigenfunction. However, the constant term contributes to the total energy of the system. This separation of electron and nuclei motion is the Born-Oppenheimer approximation.[28] The latter approach does not account for the coupling of electronic states and is therefore an adiabatic approximation. Using this approximation, we can treat a molecular structure as a configuration of fixed nuclei $\boldsymbol{R} = \{\boldsymbol{R}_A\}$ and calculate its electronic energy. By performing these calculations for every possible $\boldsymbol{R}$, we can construct a PES that describes the energy of the system E as a function of the nuclear positions for a given electronic state. To determine the energy of the system, we take the expectation value of the Hamiltonian with its normalized ground state wave function:

$$E = \langle \Psi(\boldsymbol{\tau}) | \hat{H}(\boldsymbol{\tau}) | \Psi(\boldsymbol{\tau}) \rangle. \tag{2.6}$$

In many cases, the exact wave function of a system is unknown, but the variational principle can be applied to obtain the best approximation to the ground state wave function. The principle states that the energy E_trial of a normalized trial wave function $\Phi(\boldsymbol{\tau})$ is always greater than or equal to the true energy E_0 of the ground state wave function:

$$E_0 \leq E_{\mathrm{trial}} = \langle \Phi(\boldsymbol{\tau}) | \, H(\boldsymbol{\tau}) \, | \Phi(\boldsymbol{\tau}) \rangle \, . \tag{2.7}$$

Therefore, the energy of the trial wave function can be used to assess the quality of the approximation by the trial wave function.

2.1.2 Density Functional Theory

Density functional theory is a widely used computational method in quantum chemistry and condensed matter physics, thanks to its ability to accurately predict the electronic properties of materials while significantly reducing computational cost compared to traditional wave function-based methods. In traditional wave function-based methods, each electron n in a system is described by its 3 spatial coordinates $\mathbf{r}$ and its spin state s, leading to $4n$ degrees of freedom for the wave function. This can result in computationally demanding calculations for systems with more than a few atoms.

DFT is based on the fundamental concept of the electron density, which is a function of the electron positions $\mathbf{r}$ and does not depend on spin states. Instead of explicitly computing the many-electron wave function, DFT calculates the total energy of a system as a function of its electron density $\rho(\mathbf{r})$. The electron density is the square of the wave function, integrated over all the electron coordinates except those of one electron and multiplied by the number of electrons n:

$$\rho(\boldsymbol{r}) = n \int ... \int |\Psi(\boldsymbol{\tau})|^2 \mathrm{d}s_1 \mathrm{d}\boldsymbol{\tau}_2 ... \boldsymbol{\tau}_n . \tag{2.8}$$

This density-based approach allows DFT to bypass the computationally expensive step of explicitly computing the many-electron wave function, significantly reducing the computational cost of electronic structure calculations while still providing accurate results.

For this to work, it is necessary for the electron density to determine the complete system. The number of electrons in the system is equal to the sum of $\rho(\boldsymbol{r})$ over the full space. The electron density exhibits cusps at the locations of the nuclei, with the height of the cusp depending on the charge of the nuclei, thereby providing information about the element. The total charge of the system can be obtained by adding the nuclear charges and the number of electrons. According to the first Hohenberg-Kohn theorem,[29] the ground state electron density satisfies this condition. If the electron density does not uniquely specify the Hamiltonian, there would exist two distinct external potentials V_{eN} and V'_{eN} that lead to the same ground state electron density $\rho_0(\boldsymbol{r})$. This would imply that there are two different Hamiltonians H and H', which would result in two different ground state wave functions Ψ_0 and Ψ'_0 and energies E_0 and E'_0. The expectation value of one Hamiltonian can be calculated with the wave function belonging to the other

$$\begin{aligned} E_0' &< \langle\Psi_0|\, H' \,|\Psi_0\rangle \ , \\ &< \langle\Psi_0|\, H \,|\Psi_0\rangle + \langle\Psi_0|\, H' - H \,|\Psi_0\rangle \ , \\ &< E_0 + \int (V_{eN}' - V_{eN})\rho_0(\boldsymbol{r})\mathrm{d}\boldsymbol{r} \ , \end{aligned} \tag{2.9}$$

and with interchanged primed and unprimed quantities,

$$E_0 < E_0' - \int (V_{\mathrm{eN}} - V_{\mathrm{eN}}')\rho_0(\boldsymbol{r})\mathrm{d}\boldsymbol{r} \ . \tag{2.10}$$

Adding equation 2.9 to equation 2.10 leads to the contradiction

$$E_0' + E_0 < E_0 + E_0' \ . \tag{2.11}$$

As shown above, the initial assumption that the electron density does not uniquely describe the system is incorrect. The second Hohenberg-Kohn theorem[29] proves that the ground state electron density can be found using the variational principle. Since the exact ground state wave function Ψ_0 corresponds to the exact ground state electron density ρ_0, the exact ground state energy E_0 can be calculated via the electron density as $E[\rho_0]$. That is,

$$E_0 = \langle\Psi_0|\, H \,|\Psi_0\rangle = E[\rho_0] \ . \tag{2.12}$$

The use of an approximate wave function $\widetilde{\Psi}_0$ and electron density $\widetilde{\rho}_0$ results in an approximate ground state energy $\widetilde{E}_0$, given by

$$\widetilde{E}_0 = \langle\widetilde{\Psi}_0|\, H \,|\widetilde{\Psi}_0\rangle = E[\widetilde{\rho}_0] \ . \tag{2.13}$$

The variational principle for wave functions dictates that E_0 must be lower than $\widetilde{E}_0$, which in turn means that $E[\rho_0]$ must be lower than $E[\widetilde{\rho}_0]$. Therefore, the variational principle can be used to approximate the ground state electron density.

According to the two Hohenberg-Kohn theorems, the ground state energy of a system can be calculated using the electron density, given by

$$E[\rho_0] = T[\rho_0] + E_{\mathrm{ee}}[\rho_0] + E_{\mathrm{eN}}[\rho_0] + V_{\mathrm{NN}} \ . \tag{2.14}$$

The terms in equation 2.14 correspond to the terms in the Hamiltonian equation 2.4. To calculate the energy $E[\rho_0]$, we need to define expressions for the kinetic energy of the electrons $T[\rho_0]$, the electron-electron repulsion $E_{\mathrm{ee}}[\rho_0]$, and the electron-nuclei interaction $E_{\mathrm{eN}}[\rho_0]$. First the nuclear-electron interaction $E_{\mathrm{eN}}[\rho_0]$ is described by the expression,

$$E_{\mathrm{eN}}[\rho_0] = \int V_{\mathrm{Ne}}\rho_0(\boldsymbol{r})\mathrm{d}\boldsymbol{r} = -\sum_{A=1}^{M}\int \frac{Z_A}{r_{iA}}\rho_0(\boldsymbol{r})\mathrm{d}\boldsymbol{r} \ . \tag{2.15}$$

The electron-electron interaction $E_ee[\rho_0]$ is separated into two parts: the classical Coulomb part, $J[\rho_0]$, and the non-classical part, $E_{\mathrm{ncl}}[\rho_0]$, which includes the exchange, correlation, and interaction corrections. The Coulomb interaction can be described classically by integrating over the complete space and adding a factor of $\frac{1}{2}$ to avoid double counting:

$$J[\rho_0] = \frac{1}{2}\int\int \frac{\rho(\boldsymbol{r}_1)\rho(\boldsymbol{r}_2)}{\boldsymbol{r}_{12}}\mathrm{d}\boldsymbol{r}_1\mathrm{d}\boldsymbol{r}_2 \ . \tag{2.16}$$

Equation 2.16 allows for an exact calculation of a large contribution of the Coulomb energy.

To describe the kinetic energy, Thomas and Fermi[30,31] first attempted to use the kinetic energy of a homogeneous electron gas, which yielded poor results. To improve the kinetic energy, Kohn and Sham introduced the idea of using one-electron orbitals, or wave functions, to split the kinetic energy functional into two parts: a part describing a non-interacting reference system and a part describing the Coulomb interaction between electrons. The reference system consists of electrons that behave like uncharged fermions, and the missing interaction between electrons due to Coulomb interaction is described by the effective potential $V_{\mathrm{S}}(\mathbf{r})$.

To calculate the kinetic energy of the non-interacting system, one-electron orbitals ϕ_i are introduced, for which the exact wave function is known, given by a Slater determinant. This Slater determinant $\Phi_{\mathrm{S}}(\boldsymbol{\tau})$ is given by a matrix of the one-electron orbitals:

$$\Phi_{\mathrm{S}}(\boldsymbol{\tau}) = \frac{1}{\sqrt{n!}}\begin{vmatrix} \phi_1(\boldsymbol{\tau}_1) & \phi_2(\boldsymbol{\tau}_1) & \cdots & \phi_n(\boldsymbol{\tau}_1) \\ \phi_1(\boldsymbol{\tau}_2) & \phi_2(\boldsymbol{\tau}_2) & \cdots & \phi_n(\boldsymbol{\tau}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \phi_1(\boldsymbol{\tau}_n) & \phi_2(\boldsymbol{\tau}_n) & \cdots & \phi_n(\boldsymbol{\tau}_n) \end{vmatrix} \ . \tag{2.17}$$

For this system the exact kinetic energy is

$$T_{\mathrm{S}}[\rho_{\mathrm{S}}] = -\frac{1}{2}\sum_{i=1}^{n}\langle\phi_i(\boldsymbol{r})|\,\nabla^2\,|\phi_i(\boldsymbol{r})\rangle \ . \tag{2.18}$$

In the Kohn-Sham approach, the Hamiltonian of the non-interacting system is the sum of one-electron operators, similar to Hartree-Fock theory.[32,33] These operators, called Kohn-Sham operators, are given by:

$$\hat{f}_{\mathrm{KS}} = -\frac{1}{2}\nabla^2 + V_{\mathrm{S}}(\boldsymbol{r}) \ . \tag{2.19}$$

The Kohn-Sham orbitals are then determined by solving the Kohn-Sham equations:

$$\hat{f}_{\mathrm{KS}}\phi_i = \epsilon_i \phi_i \ , \tag{2.20}$$

where ϵ_i is the energy of the i-th orbital. The effective potential $V_{\mathrm{S}}(\boldsymbol{r})$ is chosen such that the electron density of these orbitals, $\rho_{\mathrm{S}}(\boldsymbol{r})$, is equal to the real density $\rho_0(\boldsymbol{r})$:

$$\rho_{\mathrm{S}}(\boldsymbol{r}) = \sum_{i=1}^{n} |\phi_i(\boldsymbol{\tau})|^2 = \rho_0(\boldsymbol{r}) \ . \tag{2.21}$$

By solving the Kohn-Sham equations self-consistently, the Kohn-Sham orbitals and the effective potential are obtained iteratively until convergence is achieved. The non-interacting reference system only accounts for a part of the kinetic energy, leaving out the part that is not described by the Coulomb interaction of the electrons. To account for this difference, we combine the non-classical part of the electron-electron interaction $E_{\mathrm{ncl}}[\rho_0]$ with the difference between the exact and reference kinetic energy to obtain the exchange-correlation functional $E_{\mathrm{XC}}[\rho_0]$:

$$E_{\mathrm{XC}}[\rho_0] = T[\rho_0] - T_{\mathrm{S}}[\rho_0] + E_{\mathrm{ncl}}[\rho_0] \ , \tag{2.22}$$

Here, $T[\rho_0]$ represents the exact kinetic energy, and $T_{\mathrm{S}}[\rho_0]$ is the kinetic energy of the non-interacting reference system. The exchange-correlation functional $E_{\mathrm{XC}}[\rho_0]$ contains both the kinetic correlation energy and the potential correlation and exchange energy. The equation to calculate the energy of the ground state electron density $E[\rho_0$ is:

$$E[\rho_0] = T_{\mathrm{S}}[\rho_0] + E_{\mathrm{eN}}[\rho_0] + J[\rho_0] + E_{\mathrm{XC}}[\rho_0] \ , \tag{2.23}$$

where $E_{\mathrm{eN}}[\rho_0]$ represents the electron-nucleus interaction energy and $J[\rho_0]$ is the classical electron-electron interaction energy.
To ensure that the effective potential $V_{\mathrm{S}}(\boldsymbol{r})$ yields Kohn-Sham orbitals that reproduce the same density as the real system, as much of the energy expression as possible must be expressed in terms of the Kohn-Sham orbitals. Therefore, the equation can be written as:

$$
\begin{aligned}
E[\rho_0] = &-\frac{1}{2}\sum_{i=1}^{n}\langle\phi_i(\boldsymbol{r})|\,\nabla^2\,|\phi_i(\boldsymbol{r})\rangle - \sum_{i=1}^{n}\sum_{A=1}^{M}\int\frac{Z_A}{r_{A1}}|\phi_1(\boldsymbol{r}_1)|^2\mathrm{d}\boldsymbol{r}_1 \\
&+\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\int\int|\phi_1(\boldsymbol{r}_1)|^2\frac{1}{r_{12}}|\phi_1(\boldsymbol{r}_2)|^2\mathrm{d}\boldsymbol{r}_1\mathrm{d}\boldsymbol{r}_2 + E_{\mathrm{XC}}[\rho_0(\boldsymbol{r})]\ .
\end{aligned}
\tag{2.24}
$$

Here, $\phi_i(\boldsymbol{r})$ represents the Kohn-Sham orbitals, and Z_A is the charge of the Ath nucleus. The first term represents the kinetic energy, the second term is the electron-nucleus interaction energy, the third term is the classical electron-electron interaction energy, and the fourth term is the exchange-correlation energy $E_{\mathrm{XC}}[\rho_0(\boldsymbol{r})]$ evaluated at the ground state electron density $\rho_0(\boldsymbol{r})$. To obtain the orbitals one can minimize the energy expression under the constraint of orthonormality using the variational principle. This results in a Euler-Lagrange equation which gives the one-electron Kohn-Sham equations:

$$
\left(-\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{N}\frac{Z_A}{r_{A1}} + \sum_{j=1}^{n}\int\frac{|\phi_j(\boldsymbol{r}_2)|^2}{r_{12}}\mathrm{d}\boldsymbol{r}_2 + \frac{\partial E_{\mathrm{XC}}[\rho_0(\boldsymbol{r})]}{\partial\rho_0(\boldsymbol{r})}\right)\phi_i = \epsilon_i\phi_i\ . \tag{2.25}
$$

These equations are expressed as a Schrödinger-like equation with an effective potential that includes the Coulomb interaction between electrons, as well as the exchange-correlation potential. The Kohn-Sham orbitals are eigenfunctions of this effective potential, and thus, the equations have to be solved self-consistently. To solve the equations, a basis set expansion is used, where the orbitals are expressed as a linear combination of basis functions:

$$
\phi_i = \sum_{\mu} c_{\mu i}\chi_\mu\ , \tag{2.26}
$$

Examples of basis functions include plane waves, Gaussians, or numerical atomic orbitals. The Kohn-Sham equations (2.25) can then be solved by transforming the differential equations into a generalized eigenvalue problem that can be solved using linear algebra.

The exchange-correlation functional E_{XC} serves as a catch-all for any unknown energy contributions in density functional theory. The accuracy of DFT calculations is heavily reliant on finding a suitable approximation for this functional. However, improving the functional is not a straightforward task. Different approaches can be organized in the Jacob's ladder of density functionals, which reflects their trend of increasing accuracy.[34] The first step on this ladder is the local density approximation (LDA), which models the exchange-correlation energy as that of a locally homogeneous electron gas. The generalized gradient approximation (GGA) builds on the LDA approach by considering not only the electron density, but also its local gradient, to account for inhomogeneity in the electron density. The meta-GGA approach goes a step further by incorporating the kinetic energy density alongside the gradient. In hybrid functionals, a fraction of exact Hartree-Fock exchange is included.

2.1.3 Spin-Polarised Calculations

To accurately describe hydrogen atoms, a closed shell approach is not sufficient. Instead, the different spin states of the electrons must be considered separately. This is achieved by defining the electron density with a spin of $s = +\frac{1}{2}$ and $s = -\frac{1}{2}$ separately, denoted as $\rho_\alpha(\boldsymbol{r})$ and $\rho_\beta(\boldsymbol{r})$, respectively. These densities are obtained by integrating the square of the wave function $|\Psi(\boldsymbol{\tau})|^2$ over all the spatial coordinates except for the first electron's coordinates, for which the spin is fixed.

$$\rho_\alpha(\boldsymbol{r}) = n \int \dots \int |\Psi(\boldsymbol{\tau})|^2 \mathrm{d}\boldsymbol{\tau}_2 \dots \mathrm{d}\boldsymbol{\tau}_n \quad \text{for } s_1 = +\frac{1}{2} \ , \tag{2.27}$$

$$\rho_\beta(\boldsymbol{r}) = n \int \dots \int |\Psi(\boldsymbol{\tau})|^2 \mathrm{d}\boldsymbol{\tau}_2 \dots \mathrm{d}\boldsymbol{\tau}_n \quad \text{for } s_1 = -\frac{1}{2} \ . \tag{2.28}$$

The total electron density $\rho(\boldsymbol{r})$ is then obtained as the sum of $\rho_\alpha(\boldsymbol{r})$ and $\rho_\beta(\boldsymbol{r})$:

$$\rho(\boldsymbol{r}) = \rho_\alpha(\boldsymbol{r}) + \rho_\beta(\boldsymbol{r}) \ , \tag{2.29}$$

The magnetization density $Q(\boldsymbol{r})$ is defined as the difference between $\rho_\alpha(\boldsymbol{r})$ and $\rho_\beta(\boldsymbol{r})$:

$$Q(\boldsymbol{r}) = \rho_\alpha(\boldsymbol{r}) - \rho_\beta(\boldsymbol{r}) \ . \tag{2.30}$$

The total ground state energy can be expressed as a functional of either $\rho_\alpha(\boldsymbol{r})$ and $\rho_\beta(\boldsymbol{r})$ or $\rho(\boldsymbol{r})$ and $Q(\boldsymbol{r})$ in a collinear spin-polarized DFT calculation, with:

$$E[\rho_\alpha, \rho_\beta] = E[\rho, Q] \ . \tag{2.31}$$

To represent both electron density and spin density, a non-interacting reference system in the Kohn-Sham approach must use separate Kohn-Sham operators for α and β electrons. This is achieved through the following operators:

$$\hat{f}^\alpha_{\mathrm{KS}} = -\frac{1}{2}\nabla^2 + V^\alpha_{\mathrm{S}}(\boldsymbol{r}) \ , \tag{2.32}$$

$$\hat{f}^\beta_{\mathrm{KS}} = -\frac{1}{2}\nabla^2 + V^\beta_{\mathrm{S}}(\boldsymbol{r}) \ . \tag{2.33}$$

While a single Slater determinant can still represent the system, there may be different spatial orbitals for α and β electrons, and the α- and β-orbitals may not be orthogonal to each other. In contrast to the restricted case, the total energy functional in spin polarised density functional theory depends on both α- and β-densities:

$$E[\rho_\alpha, \rho_\beta] = T_\mathrm{S}^\mathrm{u}[\rho_\alpha, \rho_\beta] + E_\mathrm{eN}[\rho_0] + J[\rho_0] + E_\mathrm{XC}^\mathrm{u}[\rho_\alpha, \rho_\beta] \ . \tag{2.34}$$

The kinetic and exchange-correlation functional components of the total energy functional also depend on the α- and β-densities, and the exchange-correlation potential for the α- and β-electrons can be obtained by deriving the exchange-correlation functional with respect to the α- and β-densities, respectively.[35] When $\rho_\alpha = \rho_\beta$, the spin density becomes zero and $T_\mathrm{S}^\mathrm{u}[\rho\alpha, \rho_\beta]$ is defined the same way as $T_\mathrm{S}[\rho]$.

2.1.4 Exchange Correlation Functionals

PBE

The PBE density functional theory functional, named after Perdew, Burke, and Ernzerhof, is a widely used GGA functional in solid and surface chemistry.[36] It is a relatively simple functional that incorporates all of its parameters from fundamental constants, which makes it useful for a wide range of systems. The functional is composed of exchange and correlation terms, each containing a LDA term $\epsilon_\mathrm{X}^\mathrm{LDA}$ and $\epsilon_\mathrm{C}^\mathrm{LDA}$, respectively, as well as a term that depends on the gradient $F_\mathrm{X}(\rho, Q, |\nabla\rho|)$ and $H(\rho, Q, |\nabla\rho|)$. The exchange-correlation functional is computed as an additive combination of these terms, given by the equation:

$$\begin{aligned} E_\mathrm{XC}^\mathrm{PBE}[\rho, Q] =& E_\mathrm{X}^\mathrm{PBE}[\rho, Q] + E_\mathrm{C}^\mathrm{PBE}[\rho, Q] \\ =& \int \rho\epsilon_\mathrm{X}^\mathrm{LDA}(\rho) F_\mathrm{X}(\rho, Q, |\nabla\rho|)\mathrm{d}\boldsymbol{r} + \int \rho \left[\epsilon_\mathrm{C}^\mathrm{LDA} + H(\rho, Q, |\nabla\rho|)\right] \mathrm{d}\boldsymbol{r} \ . \end{aligned} \tag{2.35}$$

The PBE functional's reliance on fundamental constants and its ability to perform well for a wide range of systems make it an attractive option for computational studies in solid and surface chemistry.

RPBE

The RPBE density functional theory functional, developed by Hammer et al.,[37] is a revised version of the PBE functional. While the two functionals are similar, the RPBE functional uses a different mathematical form for the exchange energy enhancement factor $F_\mathrm{X}(\rho, Q, |\nabla\rho|)$ compared to the PBE functional. This change has a significant effect on the performance of the functional, reducing the overbinding of chemisorbed atoms and molecules by approximately half compared to the PBE functional. This improvement makes the RPBE functional particularly useful for studying systems where chemisorption plays a critical role, such as surface chemistry or heterogeneous catalysis. Overall, the RPBE functional represents a valuable advancement in the field of density functional theory and has found widespread use in the study of materials and chemical systems.

PBE0

The PBE0 density functional theory functional is a hybrid GGA functional that combines the exchange energy from both the PBE functional and Hartree-Fock theory. This functional was developed by Adamo and Barone[38] and provides an improved description of electronic structures for a wide range of molecular and solid-state systems. The PBE0 functional has proven to be a powerful tool in the study of chemical reactions, excited states, and molecular properties. The inclusion of Hartree-Fock exchange energy improves the accuracy of the functional, especially for systems with strong correlation effects. The Hartree-Fock exchange is simply added to the exchange-correlation energy:

$$E_{\mathrm{XC}}^{\mathrm{PBE0}} = E_{\mathrm{XC}}^{\mathrm{PBE}} + \frac{1}{4}\left(E_{\mathrm{X}}^{\mathrm{HF}} - E_{\mathrm{X}}^{\mathrm{PBE}}\right) \tag{2.36}$$

The PBE0 functional also performs well for transition metal-containing systems, which can be difficult to describe accurately using standard GGA functionals.

2.1.5 Periodic Boundary Conditions

To construct a realistic surface for interaction with gas phase particles, a large number of atoms are required to create an uninterrupted surface. However, simply extending the surface by adding more atoms, to circumvent creating artificial interruptions in the surface, increases the cost of the calculations immensely. This is where periodic boundary conditions (PBC) come into play.

The idea behind PBC is to place the atoms generating the surface into a box that is replicated in all directions. This results in a slab of atoms that is infinite in two directions and is separated by vacuum in the third direction. In this direction, the slab is continuously repeated. If an atom in the slab moves out of the box in one direction, its image will enter from the opposite direction, ensuring that the surface remains uninterrupted.

To apply periodic boundary conditions to the system, a periodic potential is required. This is where Bloch's theorem[39] comes in. Bloch's theorem states that any eigenfunction $\phi_{\boldsymbol{k}}^{n}(\boldsymbol{r})$ can be expressed as a product of a function $u_{\boldsymbol{k}}^{n}(\boldsymbol{r})$ with the periodicity of the lattice and a plane wave $\mathrm{e}^{\boldsymbol{ikr}}$, where n is the band index and $\boldsymbol{k}$ is a vector in the first Brillouin zone:[40]

$$\phi_{\boldsymbol{k}}^{n}(\boldsymbol{r}) = u_{\boldsymbol{k}}^{n}(\boldsymbol{r})\mathrm{e}^{n\boldsymbol{kr}} \ . \tag{2.37}$$

The Brillouin zone is the Wigner-Seitz[41] cell of the reciprocal space. The two parts of the eigenfunction, $u_{\boldsymbol{k}}^{n}(\boldsymbol{r})$ and $\mathrm{e}^{\boldsymbol{ikr}}$, can be considered separately. The first part must be a function with the periodicity of the lattice, for which plane waves are a natural choice. In addition to using plane waves to describe the periodicity of the lattice, it is also possible to describe it by periodically replicating local basis functions χ_i^0 as

$$\chi_i^n(\boldsymbol{r} - \boldsymbol{R}_n) = \chi_i^0(\boldsymbol{r}) \ , \tag{2.38}$$

with

$$\boldsymbol{R_n} = n_1\boldsymbol{a}_1 + n_2\boldsymbol{a}_2 + n_3\boldsymbol{a}_3 \ , \tag{2.39}$$

where $\boldsymbol{a}_{1,2,3}$ are the lattice vectors of the periodic box and $n_{1,2,3}$ are integer numbers denoting the position of the periodic image. This approach has the advantage that consistent basis functions can be used even without periodic boundary conditions.

The second part of Bloch's theorem contains the vector $\boldsymbol{k}$, which can be any vector in the first Brillouin zone. However, since the eigenvalues depend on the $\boldsymbol{k}$-vectors, also called $\boldsymbol{k}$-points, the selection of $\boldsymbol{k}$-points is crucial. Using more $\boldsymbol{k}$-points leads to a more accurate description of the system, but for each selected $\boldsymbol{k}$-point, the Kohn-Sham equations must be solved separately. Therefore, the minimal number of $\boldsymbol{k}$-points needed to achieve convergence should be chosen.

From the chosen $\boldsymbol{k}$-points, a weighted sum of energies is calculated. This means that the $\boldsymbol{k}$-point grid must be checked to find the point at which increasing the $\boldsymbol{k}$-point density yields energy convergence to the desired accuracy.

2.1.6 Dispersion Correction

The accurate description of long-range van der Waals (vdW) interactions between atoms and molecules is crucial in many areas of chemistry and materials science. While standard DFT accounts these interactions only incompletely, the addition of a pairwise interatomic C_6R^{-6} term to the DFT energy has been a popular approach[42,43] to include vdW interactions in DFT calculations, yielding the DFT+vdW method:

$$E^{\text{DFT+vdW}} = E^{\text{DFT}} + E_{\text{vdW}} \tag{2.40}$$

Here, E^{DFT} is the DFT energy and E_{vdW} is the vdW energy correction, which takes the form:

$$E_{\text{vdW}} = -\frac{1}{2}\sum_{A,B} f_{\text{damp}}(R_{AB}, R_A^0, R_B^0) C_{6AB} R_{AB}^{-6} \ , \tag{2.41}$$

where C_{6AB} and R_{AB}^{-6} are the C_6 coefficient and the distance-dependent R^{-6} term, respectively, for the interaction between atoms A and B, and the damping function

$f_{\mathrm{damp}}(R_{AB}, R_A^0, R_B^0)$ is a function of the vdW radii of the atoms and the interatomic distance R_{AB}.

The Tkatchenko-Scheffler (TS) correction[44] is a widely used method to calculate the vdW interaction energies, where the C_6 coefficient and the vdW radii are determined from the ground state electron density of DFT or Hartree-Fock theory and reference values of the free atoms using the Hirshfeld partitioning of the electron density.[45] In this method, the damping function $f_{\mathrm{damp}}(R_{AB}, R_{AB}^0)$ is a function of the sum of the vdW radii R_{AB}^0, given by:

$$f_{\mathrm{damp}}(R_{AB}, R_{AB}^0) = \left[1 + \mathrm{e}^{-d\left(\frac{R_{AB}}{s_R R_{AB}^0} - 1\right)}\right]^{-1} . \tag{2.42}$$

where d is a parameter that adjusts the steepness of the damping function, and s_R is a free parameter that is fitted to a database of converged Coupled Cluster Singles and Doubles with perturbative triples [CCSD(T)] binding energies of 22 different dimers with varying interaction strengths. The value of d is typically set to 20, providing a satisfactory description of binding energy curves for rare gas and vdW-bonded organic molecule dimers.

2.2 Machine Learning Potentials

2.2.1 Introduction to Machine Learning Potentials

While first principles methods like DFT offer accurate predictive power and agree well with experimental data for various systems, their application is limited to smaller systems due to their high computational costs. Machine learning potentials that are trained on electronic structure data can maintain the accuracy of first principle methods while speeding up energy and force evaluations by several orders of magnitude. This makes it possible to extend the length, time and quantity scales of first principles-quality atomistic simulations to more complex systems. Several studies[6,46–49] have demonstrated the ability of MLPs in accelerating the simulation of complex systems while maintaining accuracy.

The extension of MLPs to larger and more complex systems has been a focus of research in recent years, and this development can be categorized into four generations.[50] The first generation of MLPs, which dates back to 1995, was developed for low-dimensional systems and relied on neural network potentials.[51–53] The second generation of MLPs was introduced in 2007 in the form of high-dimensional neural network potentials.[21,50,54–56] HDNNPs enable the simulation of high-dimensional systems containing tens of thousands of atoms by approximating the Major part of the atomic interactions to be local. The energy is expressed as a sum of atomic energy contributions. These atomic energy contributions are estimated based on the local structrual environment which is represented within a cutoff sphere that is translationally, rotationally, and permutationally invariant. This generation of MLPs has been widely used due to its cost-effectiveness and applicability to a broad range of systems. Examples of second-generation MLPs include Gaussian approximation potentials,[57] moment tensor potentials,[58] and many others.[59–61] A variety of structural descriptors have also been developed for second-generation MLPs, such as atom-centered symmetry functions (ACSF),[62] bispectrum,[57] Coulomb matrices,[63] overlapp matrix,[64] smooth overlap of atomic positions,[65] and many others.[66]

Recent developments in machine learning potentials have resulted in the third and fourth generations of MLPs. The third generation MLPs take into account the long-range interactions beyond the cutoff sphere, such as electrostatics and dispersion. To include the electrostatic interaction, element-specific fixed charges[57,67] or local environment-dependent atomic charges can be used.[68–70] The fourth generation MLPs further incorporate non-local and global dependencies, allowing for the description of non-local charge transfer and different global charge states. This generation includes methods such as charge equilibration neural networks,[71] Becke population neural networks,[72] and fourth-generation HDNNPs.[73,74]

2.2.2 High Dimensional Neural Network Potentials

One of the most used MLPs is the second generation HDNNP,[21] which employs a separate feed-forward neural network potential for each type of element included in a structure. The energy contribution E_n^m of each element M included N_m times in the structure is calculated via these neural networks, and the energy of the entire system E_{tot} is obtained by summing up the energy values for each atom:

$$E_{\text{tot}} = \sum_{m=1}^{M} \sum_{n=1}^{N_m} E_n^m \ . \tag{2.43}$$

This approach works well for up to four different elements, but adding more elements exponentially increases the configurational space. Feed-forward neural networks were the first and simplest type of artificial neural networks used for neural network potentials, they can only calculate the energy of one structural setup since any change in the number or ordering of atoms changes the number or input of the input nodes in the neural network. However, feed-forward neural networks are effective in calculating the energy of individual atoms. Figure 2.1 illustrates the structure of a feed-forward neural network.

A feed-forward neural network comprises interconnected layers that process information in one direction and culminate in the output layer. The input layer, which utilizes

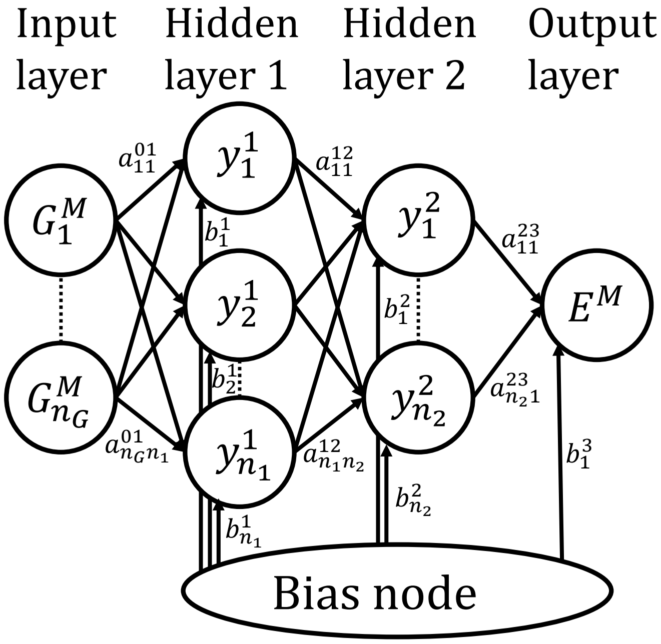


Fig. 2.1: The figure illustrates a feed forward neural network that analyzes the structural information of an atom, of element M, using structural descriptors G^M. The network consists of neurons, denoted as y, which process this information and ultimately produce the energy value E^M of the atom. The weight parameters a and b are applied to the processed values to establish the appropriate relationship between G and E. (graphic taken from [75])

symmetry functions G_i, provides data about the atomic environment to the network by characterizing the chemical environment of an atom up to a cutoff radius. The hidden layers, which usually consist of one to three layers containing up to 50 neurons each, receive and process information from the input layer before transmitting it to the output layer. Neurons in the hidden layers obtain input from every node in the preceding layer, weighted by a parameter a_{kl}^{ij}. All values are summed up by the neuron y_l^j, after which a bias weight b_l^j is added. A continuous non-linear activation function f_l^j such as a hyperbolic tangent or Sigmoid function evaluates this value in the neuron, with the bias weight being added to ensure the value lies in the non-linear region of the activation function. The output of the activation function is used as input to the following layer. The output layer uses a linear activation function to output energy without any restrictions. The atomic energy E_n^m is the sum of all weighted values of every node in the last hidden layer plus a bias weight. A separate atomic neural network is fitted for each element M. The complete equation of the feed-forward neural network with 2 hidden layers is as follows:

$$E_n^m = f_1^3 \left(b_1^3 + \sum_{l=1}^{n_2} a_{l1}^{23} \cdot f_l^2 \left(b_l^2 + \sum_{k=1}^{n_1} a_{kl}^{12} \cdot f_k^1 \left(b_k^1 + \sum_{i=1}^{n_G} a_{ik}^{01} \cdot G_i \right) \right) \right) \quad . \tag{2.44}$$

As previously mentioned, the data used to train a neural network must be invariant under translation, rotation, and permutation. Symmetry functions play a crucial role in achieving this. These functions are many-body functions centered on individual atoms, which describe the environment around them within a cutoff radius R_c. Choosing the correct value of R_c is essential, as it should encompass all relevant interactions. Typically, R_c falls between 6 and 10 Å. Importantly, since symmetry functions are local in nature, HDNNPs can model arbitrarily large systems by stitching together many different local parts.

Forces can be analytically computed from symmetry functions. Therefore the functions must be continuous and differentiable. Additionally, the effective cutoff radius for forces is twice the normal R_c. For instance, two atoms, i and j, located to each other at a distance approaching R_c, will experience forces influenced by the symmetry functions of the other atom, which can extend up to a distance of $2 \cdot R_\mathrm{c}$ away from i.

Symmetry functions are typically divided into two types: radial symmetry functions G_i^rad and angular symmetry functions G_i^ang. Radial symmetry functions depend on the distance between the central atom i and other atoms j and are defined as:

$$G_i^\mathrm{rad} = \sum_j \mathrm{e}^{-\eta (R_{ij} - R_\mathrm{S})^2} \cdot f_\mathrm{c}(R_{ij}) \tag{2.45}$$

where R_{ij} is the distance between atoms i and j, η is a scaling parameter, and R_S is a reference distance. The function $f_\mathrm{c}(R_{ij})$ is a cutoff function that smoothly decreases the contribution of atoms up to a cutoff distance R_c. Specifically, it is defined as

$$f_\mathrm{c}(R_{ij}) = \begin{cases} 0.5 \left[\cos\left(\dfrac{\pi R_{ij}}{R_\mathrm{c}} \right) + 1 \right] \text{ for } R_{ij} \leq R_\mathrm{c} \\ \hfill 0 \text{ for } R_{ij} \geq R_\mathrm{c} \end{cases} \quad . \tag{2.46}$$

The summation over all atoms ensures that the number of symmetry functions remains constant even as the number of atoms in the cutoff sphere changes during a simulation. This is necessary because the number of neurons in a neural network cannot be changed once it has been trained on a fixed set of data.

Angular symmetry functions describe the structural relation between three atoms i, j, and k and are defined as:

$$G_i^\mathrm{ang} = 2^{1-\zeta} \sum_j \sum_{k \neq j} [1 + \lambda \cos(\theta_{ijk})]^\zeta \cdot f_\mathrm{c}(R_{ij}) \cdot f_\mathrm{c}(R_{ik}) \cdot f_\mathrm{c}(R_{jk}) \ , \tag{2.47}$$

where θ_{ijk} is the angle between the vectors pointing from atom i to j and from atom i to k. The parameters ζ and λ control the shape of the angular symmetry functions. These functions depend on the parameters η, R_S, ζ, and λ, which can be optimized to accurately capture the chemical environment of the atoms.

Figure 2.2 provides examples of the symmetry functions used in this work. The total number of symmetry functions required for a system depends on the complexity of its chemical environment and the number of elements present. Typically, there are between 20 and 200 symmetry functions per atom.

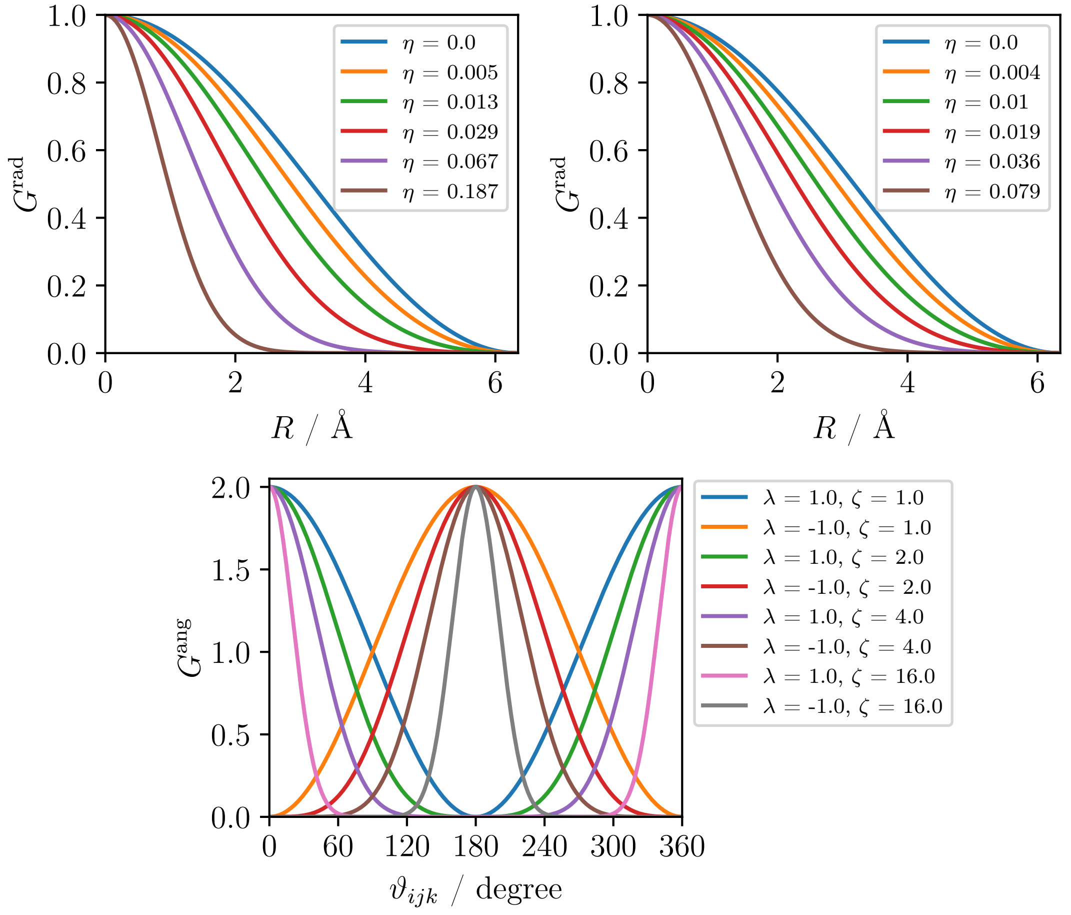


Fig. 2.2: The figure displays radial and angular symmetry functions for H, O, and Al. The top-left panel shows the radial symmetry functions for H, while the top-right panel shows the radial symmetry functions for O and Al. In both cases, the radial symmetry functions are calculated with $R_\mathrm{S} = 0\,\text{Å}$ and $R_\mathrm{c} = 6.35\,\text{Å}$. The bottom panel shows the angular symmetry functions for H, O, and Al with $\eta = 0.0\,\text{Å}^{-2}$. (graphic taken from [75])

A complete high dimensional neural network is depicted in figure 2.3. The Cartesian coordinates $\boldsymbol{R}_n^m$ of a given structure are transformed into symmetry functions $\boldsymbol{G}_n^m$,

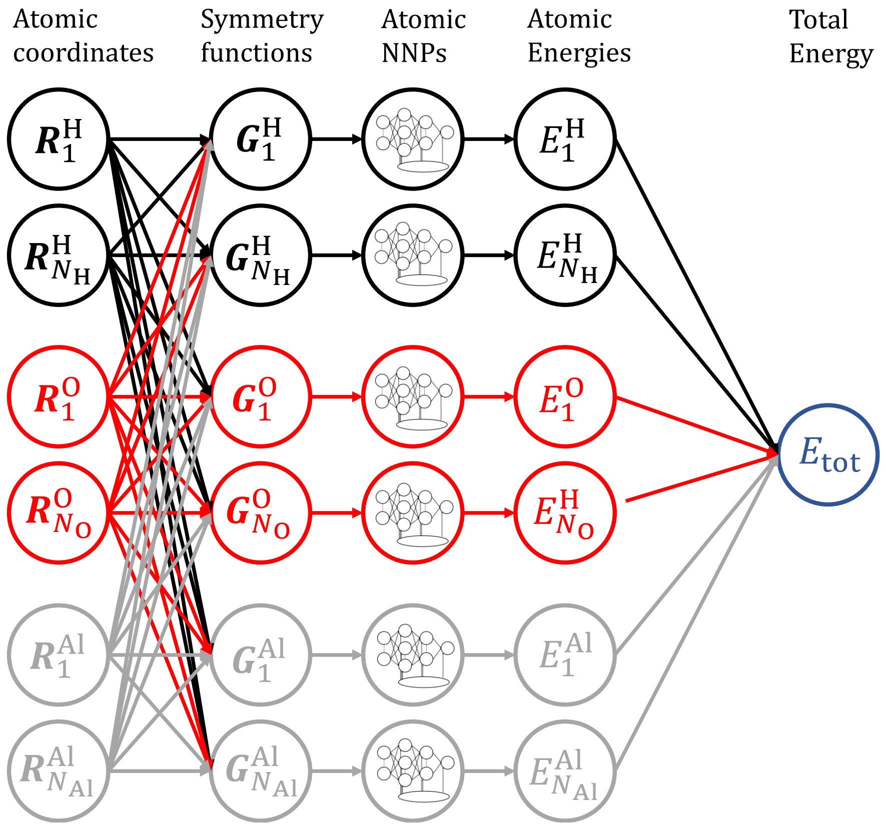


Fig. 2.3: The figure illustrates the architecture of a high-dimensional neural network potential. In this approach, the atomic coordinates $\boldsymbol{R}$ are first transformed into symmetry functions $\boldsymbol{G}$, which are then utilized by the atomic neural network potentials (NNPs) to calculate the atomic energies E. These atomic energies are subsequently summed up to obtain the total energy E^{tot}. (graphic taken from [75])

which are then processed by the atomic NNPs to generate atomic contributions E_n^m to the total energy E_{tot}. To fit the weight parameters a_{kl}^{ij} and the bias weights b_k^j to the reference data, a gradient-based iterative optimization algorithm is employed, to minimize the root mean squared error (RMSE) of the predicted total energies and forces. The global adaptive Kalman filter has been shown to perform well in this process.[76,77]
During validation, the RMSE serves as a preliminary indication of the accuracy of the fit and helps detect over- or underfitting, as well as gaps in the configurational space. A high test RMSE could result from overfitting or be caused by structures which are not well described by the data set, hinting towards gaps in the configurational space. Figure 2.4 illustrate the effects of overfitting and underfitting.
The flexibility of a neural network is determined by the number of neurons it has. As the number of neurons increases, so does the flexibility of the network. However, an excessive number of neurons can lead to overfitting. To detect overfitting, a common approach is to set aside around 10% of the reference data as a testing set and evaluate

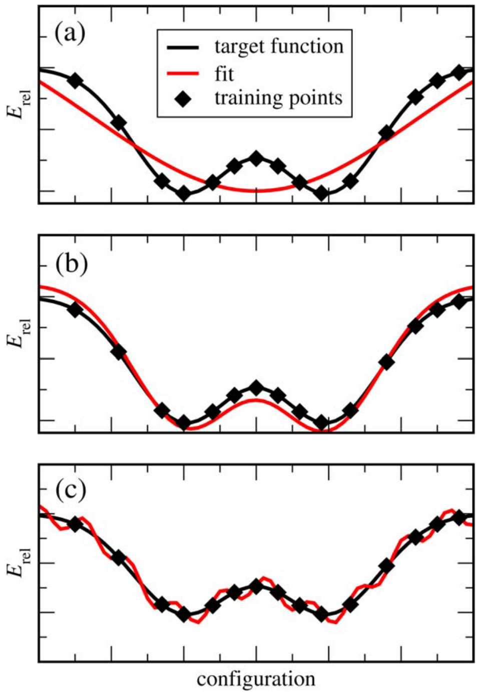


Fig. 2.4: Representation of (a) underfitting, (b) an accurate fit and (c) overfitting (graphic taken from [55]).

how accurately the neural network predicts the structures it was not trained on. If the RMSE of the test set is larger than that of the training set, it may indicate overfitting. Nevertheless, outliers, or an inhomogeneous or incoherent reference data set could also cause a similar pattern. If both the training and testing RMSEs are too large, it may be necessary to investigate the flexibility of the network and the descriptive power of the symmetry functions. In this case, adjusting the number of neurons and symmetry functions may be helpful. However, if the form of the function is not the main cause of the accuracy problem, it may be necessary to evaluate the training set. The training set should sufficiently describe the PES as the neural network's ability to extrapolate is limited. Extrapolation can be detected by checking if the value of a symmetry function is beyond the range that the neural network has been trained on. Another challenge is to identify inaccurate predictions of structures within the range of the symmetry functions. To address this issue, a second or even more neural networks with an equally good fit should be trained. If the energy or force predictions of this second neural network show a significant difference, it suggests that the PES in that region of the configuration space is not adequately represented in the training set. Structures that caused extra- or interpolation errors can be discovered by conducting MD or Monte Carlo (MC) simulations.

The following reviews [50, 54–56, 78] provide more comprehensive and detailed information about HDNNPs and ACSFs.

2.2.3 Molecular Dynamics

Following the Born Oppenheimer approximation, it is possible to simulate the classical motion of nuclei based on the PES $V(\boldsymbol{R}_A)$, which determines the forces $\boldsymbol{F}_A$ acting on the nuclei A, as given by the expression:

$$\boldsymbol{F}_A = -\frac{\partial V(\{\boldsymbol{R}_A\})}{\partial \boldsymbol{R}_A} \ . \tag{2.48}$$

To calculate the time evolution of the system, the equation of motion proposed by Newton is utilized, with the mass m_A of nucleus A being a contributing factor:

$$\boldsymbol{F}_A = m_A \ddot{\boldsymbol{R}}_A \ , \tag{2.49}$$

However, finding an analytic solution for this many-body problem is not possible. Classical MD[79,80] is utilized to find an approximate solution, which involves numerically solving the coupled differential equations with discrete time steps δt.

To numerically solve Newton's equations of motion, an algorithm that conserves energy and angular momentum, is time reversible, and symplectic is necessary. The velocity Verlet algorithm[81] satisfies these conditions, except for energy conservation due to the finite time steps. However, using sufficiently small time steps can yield satisfactory energy conservation. The algorithm propagates the positions $\boldsymbol{R}$ and velocities $\boldsymbol{v}$ in time t by first calculating the updated positions $\boldsymbol{R}(t+\delta t)$ and forces $\boldsymbol{F}(t+\delta t)$ as:

$$\boldsymbol{R}(t+\delta t) = \boldsymbol{R}(t) + \boldsymbol{v}(t)\delta t + \frac{\boldsymbol{F}(t)}{2m}\delta t^2 \ , \tag{2.50}$$

and then updating the velocities $\boldsymbol{v}(t+\delta t)$ as:

$$\boldsymbol{v}(t+\delta t) = \boldsymbol{v}(t) + \frac{\boldsymbol{F}(t) + \boldsymbol{F}(t+\delta t)}{2m}\delta t \ . \tag{2.51}$$

This numerical integration of Newton's equations of motion provides a practical and efficient method for simulating the classical motion of nuclei in the framework of the Born Oppenheimer approximation. To ensure that the energy of the system remains constant during the simulation, the microcanonical (NVE) ensemble is utilized, with N denoting the number of particles and V representing the volume of the cell. However, for a more accurate representation of a system that is in equilibrium with its environment, the canonical ensemble (NVT) should be employed, where the temperature is kept constant instead of the energy. To achieve this, a thermostat is required, such as the Nosé-Hoover Thermostat[82–84] (NH). The NH thermostat simulates a coupling to a heat bath, extending the $6N$-dimensional phase space by two artificial degrees of freedom and achieving a constant system temperature. It's worth noting, however, that the NH thermostat alone may not result in ergodic dynamics, which can be restored by utilizing a nested chain of NH thermostats.[85] By doing so, the simulated system can be effectively transformed into the canonical ensemble.

2.2.4 Experimental Setup

The theoretical results gathered in this thesis will be compared to experimental data. The experimental scattering apparatus has been described elsewhere.[18] Briefly, using photodissociation of a supersonic molecular beam of hydrogen iodide via pulses of KrF eximer laser light at 248.35 nm hydrogen atoms with incident kinetic energies $E_{\text{in}} = 1.92(2)\,\text{eV}$ and $E_{\text{in}} = 0.99(2)\,\text{eV}$ were created. The hydrogen atoms then scattered from the room temperature [0001] Al_2O_3-surface at a azimuthal incident angle $\phi_{\text{i}} = 0°$ and a polar incident angles $\theta_{\text{i}} = 40°$ and $55°$ and then were exited to a long-lived Rydberg state by two laser pulses. The first pulse excites the $1s$ -> $2p$ transition at 121.6 nm and the second one the $2p$ -> $34p$ transition at 365.9 nm. The travel time of the metastable atoms is then measured after a field-ionization and detection by a multichannel plate detector. The arrival time is then recorded by a multichannel scaler and via the calibrated flight time the velocity and kinetic energies of the H atoms can be calculated. The [0001] Al_2O_3-surface held in the UHV chamber are cleaned by annealing for several hours in oxygen (10–6 mbar) at 600–750 °C.

3 Computational Details

3.1 Density Functional Theory

3.1.1 Fritz Haber Institute-Ab Initio Materials Simulations

All DFT calculations were performed using the Fritz-Haber-Institute ab initio materials simulations (FHI-aims)) package (version 160328_3 from October 23, 2016).[86,87] The RPBE,[37] PBE[36] and PBE0[38,88] functionals were used, including the Tkatchenko-Scheffler dispersion correction.[44] For all elements, the "light" settings were used for the basis set of the numeric atom-centered basis functions. The atomic energies for H, O and Al were calculated without periodic boundary conditions, while periodic boundary conditions were applied to the bulk and slab calculations for α-Al_2O_3. Γ-centred $\boldsymbol{k}$-point grids have been used for all periodic systems. For bulk α-Al_2O_3 structures a $6{\times}6{\times}2$ $\boldsymbol{k}$-point grid was used, while the $2{\times}2$ slab super cells used a $2{\times}2{\times}1$ $\boldsymbol{k}$-point grid. For calculations of the α-Al_2O_3 slab including H atoms collinear spin polarization was applied, while for the bulk and slab structures without H atoms restricted closed-shell calculations were performed. The H atoms were initialized with a "initial_moment" of 1. Atomic spins were calculated using the Hirshfeld spin moments.[45] The calculations were converged in a self-consistent field (SCF) procedure to a value of 10^{-6} eV for energies and $10^{-4}\,\mathrm{eV\,\AA^{-1}}$ for forces. Geometry optimizations were done using the Broyden-Fletcher-Goldfarb-Shanno algorithm[89–92] with an energy convergence criterion of 0.005 eV. Convergence tests for the $\boldsymbol{k}$-point grid can be found in section 3.3.1.

3.2 Molecular Dynamics Simulations using LAMMPS

To run MD simulations, the Large-scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) (version from March 16, 2018)[93] using the n2p2 extension to include HDNNPs[94][95] was used. MD simulations for active learning were run in the ensemble NVT and NVE with a time step of $\delta t = 0.5$ fs for simulations without a H atom and $\delta t = 0.25$ fs for simulations with a H atom. The velocity Verlet algorithm was used as the integrator. To simulate the NVT ensemble a Nose Hoover thermostat[85] with a temperature damping factor of 0.01 ps was used. For the trajectories using experimental incident kinetic energies and angles, the slab structures and atom velocities were taken from the NVE trajectories at 300 K every 0.1 ps. The H atom was then initialized at a random position 6.3 Å above the surface with velocity vectors corresponding to the experimental kinetic energy, polar angle and azimuth angle. The trajectory then was run for 0.5 ps or until the H atom was at a distance of 7.8 Å from the slab.

3.2.1 Phonopy Calculations

Phonon band structures were calculated with the phonopy code (version v.2.17.1)[96,97] using both force constants from the FHI-aims RPBE DFT calculations, as well as force constants calculated using the HDNNP. Force constants were calculated using a geometry optimized primitive α-Al_2O_3 cell with 10 atoms. Geometry optimization for the HDNNP calculation was carried out using LAMMPS with the n2p2 extension. To determine the correct symmetry, the symmetry precision threshold was set to $5 \cdot 10^{-4}$. The graphs were created using the Python interface. The interface between phonopy and RuNNer was created using a Python script.

3.3 High-Dimensional Neural Network Potentials

3.3.1 Construction of the Reference Data Set

Overview

The prediction of a HDNNP is not based on physical foundations, but relies solely on the information it learned from the data set it was trained on. Because the HDNNP can represent the reference PES with high accuracy, it is important to consider the completeness, consistency, and quality of the data set when creating it. Without a complete data set, not all reaction pathways can be studied with the same accuracy, leading to a biased or inaccurate description of the system. If the data set is inconsistent, the MLP will have difficulty fitting the data, resulting in higher errors and unreliable behavior. Finally, only with high-quality calculations will the MLP be able to produce high-quality results that are free from errors or bias. To ensure these criteria of the data set, the creation of reference data is a multistep process of training, validation, and extension, commonly referred to as active learning.

During the creation of this thesis, three separate data sets have been created, using three different approaches for the initial generation of structures. The first data set had no initial reference structures, while the next two were able to start using either the structures or the potentials of their predecessors. Although the initial generation of structures was very different for all three data sets, further improvement and extension followed the same principles of active learning.

General Structure Design

In order to give an accurate description of the system, the training structures for the data set must be carefully constructed. The data set consists of three different structure types, a bulk structure, a slab structure and a structure with a slab and H atoms. In this section, convergence calculations based on the RPBE DFT functional are shown, while the PBE convergence calculations are described elsewhere.[75] The bulk structures are in the data set to ensure that the HDNNP is able to extrapolate to bulk geometries in the middle of the slab. The structure consists of one unit cell of α-Al_2O_3 , which includes 30 atoms. The calculation can introduce many bulk chemical environments into the fit, enhancing its ability to accurately describe the interior of the slab compared to slab calculations. The calculations use a 4×4×2 $\boldsymbol{k}$-point grid which was converged in energy to below 1 meV compared to a 7×7×3 $\boldsymbol{k}$-point grid. The energy convergence is shown in the top panel of figure 3.3.
For the convergence calculations for bulk structures the formation energy ΔE_{form} was used, derived from the reaction from the elements,

$$2\mathrm{Al}_{(\mathrm{s})} + \frac{3}{2}\mathrm{O}_{2(\mathrm{g})} \rightarrow \alpha\text{-}\mathrm{Al}_2\mathrm{O}_{3(\mathrm{s})} \ , \tag{3.1}$$

while for the slab calculations the energy difference between the 3 H atom positions above the unoptimised slab, shown in figure 3.2, was used. For the different amounts of layers and different super cell sizes, the relative positions of the H atom stayed the same. The three energy differences, $\Delta E_1 = E_{\mathrm{Pos1}} - E_{\mathrm{Pos2}}$, $\Delta E_2 = E_{\mathrm{Pos2}} - E_{\mathrm{Pos3}}$ and $\Delta E_3 = E_{\mathrm{Pos3}} - E_{\mathrm{Pos1}}$ were calculated from the respective energies. Slab structures must ensure an accurate description of the surface, as well as the H atom on the surface, while cost-effectively describing all the chemical environments needed for the H atoms. The $\boldsymbol{k}$-point grid for a slab created from a unit cell of α-Al_2O_3 was set to 4×4×1, with an energy convergence compared to a 10×10×2 $\boldsymbol{k}$-point grid of less than 1 meV shown in the lower panel of figure 3.3. Using a 2×2 super cell decreases the interaction energy between the H atoms in periodic images to around 1.5 meV compared to the 3×3 super cell shown in figure 3.1. Increasing the amount of O-Al-Al layer from six in the bulk unit cell to eight in the slab ensures that the symmetry functions of the H atoms on the surface will go to 0 in the middle of the slab. The energy for the number of slab layers is converged to 2 meV compared to twelve layers. More layers increase the computational time more than the small increase in energy convergence, and "bulk"-like environments would result in a significant enough increase in accuracy. The slab structure for MD simulations using experimental conditions consist of a 2×2 super cell with 8 O-Al-Al layers. The slab is terminated with the [0001] surface of α-Al_2O_3 and was relaxed with constrained bulk lattice constants in the x and y directions. The surface termination ends with one Al per unit cell on top of the O layer, which is called the half metal termination in literature.[98] A vacuum of 21 Å was added between the slabs in the z

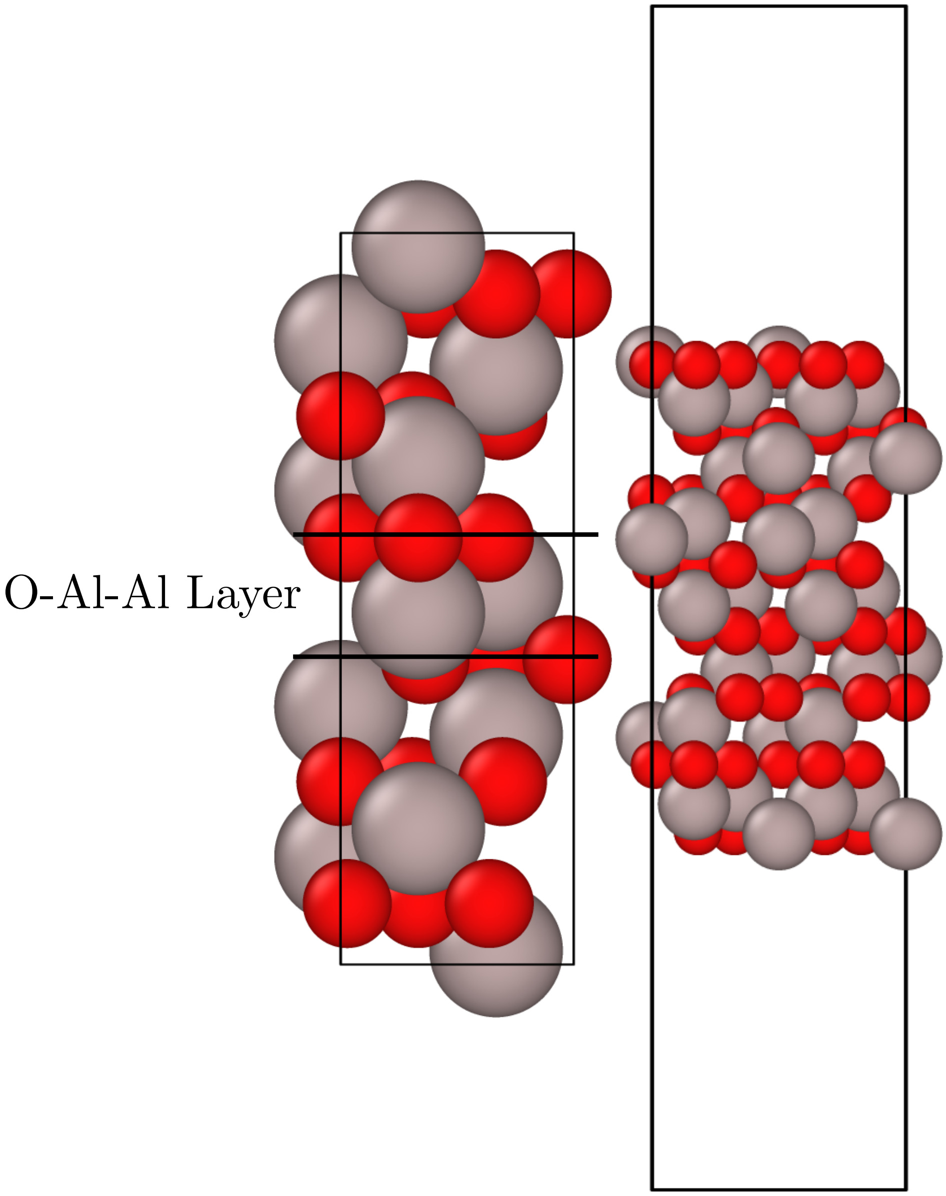


Fig. 3.1: Representation of the bulk unit cell (left) and 2×2 slab super cell (right). The red spheres are O atoms, and the gray spheres are Al atoms. One O-Al-Al layer is highlighted between the two black bars.

direction. Adding a H atom at a distance of 7 Å from the surface, with a symmetry function cutoff of $12\,a_0$ or 6.35 Å, which is a commonly used standard, gives plenty space to put a H atom at the opposite site of the slab at the same distance while keeping a distance of 7 Å between the H atoms.

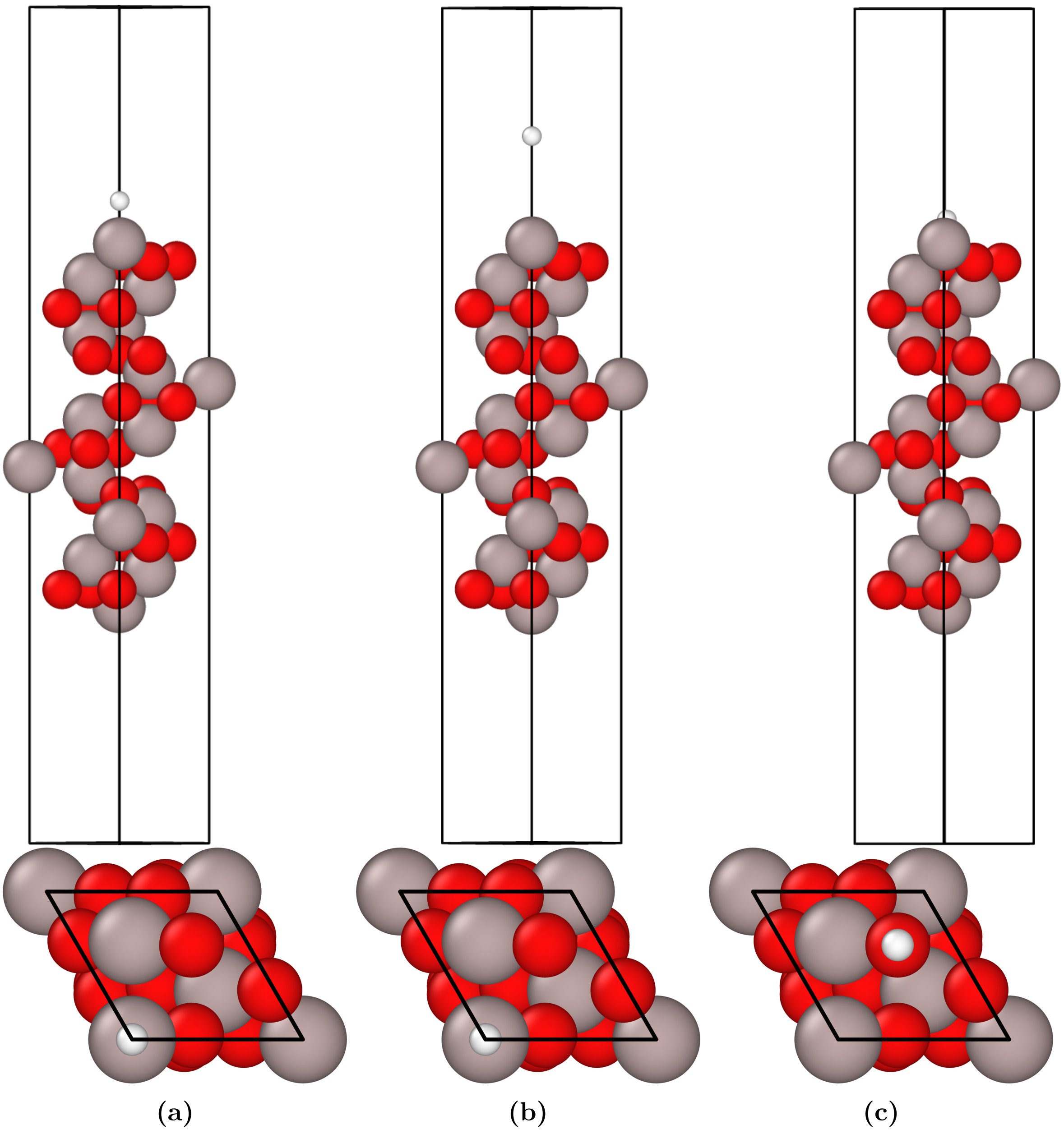


Fig. 3.2: Representation of the three different positions of the H atom. The red spheres are O atoms, the grey spheres are Al atoms and the white spheres are H atoms. (a) is the first position, (b) is the second position, and (c) is the third position. The upper part depicts the slab from the side, while the lower part shows it from the top.

Initial Structure Generation for the PBE Data Set

The initial data set for a MLP typically is based on structural data generated from experiments, for example the bulk crystal structure of α-Al_2O_3. The structure is optimized using the DFT functional to generate the first data point. Geometry-optimized

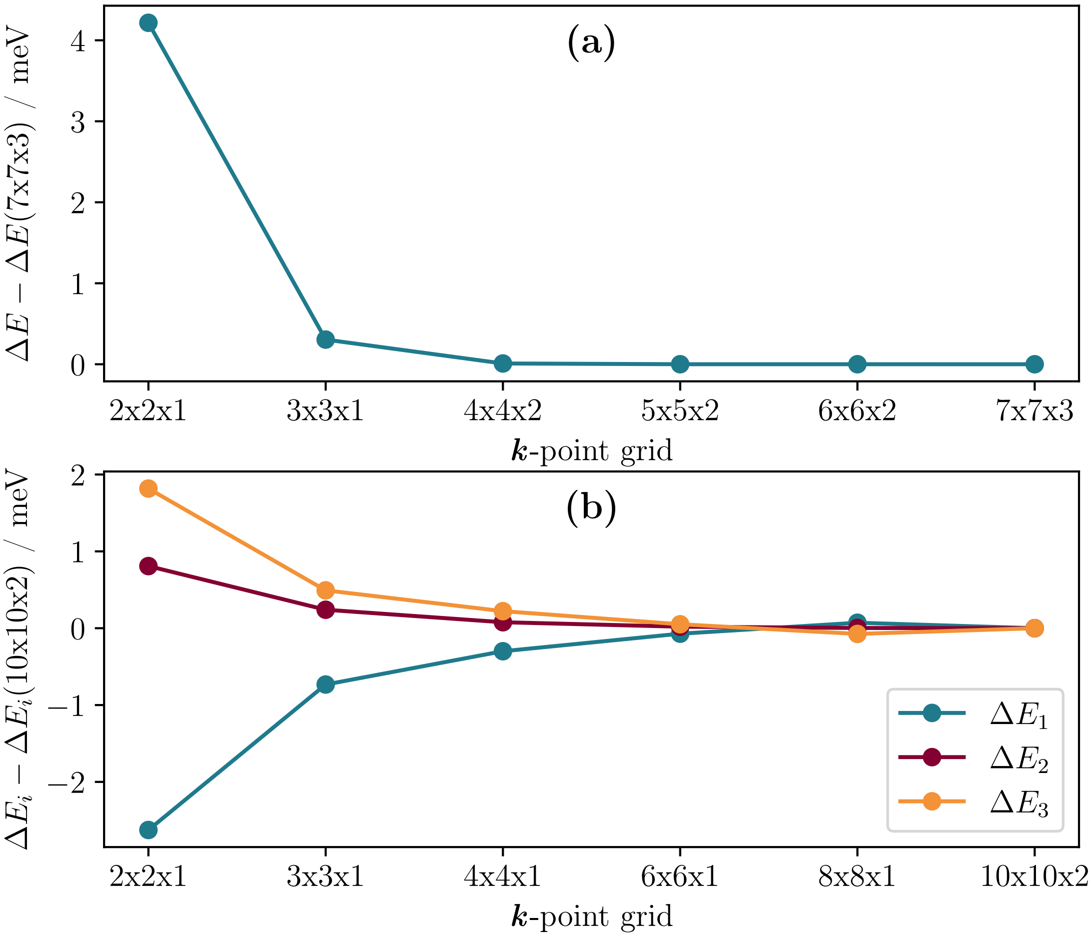


Fig. 3.3: Energy convergence for different $\boldsymbol{k}$-point grid settings. **(a)** Bulk unit cell of α-Al_2O_3 converged to a $\boldsymbol{k}$-point grid of 7×7×3. The $\boldsymbol{k}$-point grid of 4×4×2 already has an energy convergence of less than 1 meV. **(b)** Differences between three H atom positions above the α-Al_2O_3 slab with 8 O-Al-Al layers. ΔE_1, ΔE_2 and ΔE_3 are the differences in energies between the positions of H atoms shown in figure 3.2. The $\boldsymbol{k}$-point grid of 4×4×2 already has energy convergence of less than 1 meV, uses the same density as the bulk cell and is scalable for the 2×2 super cell.

structures are an ideal starting point for the construction of a reference data set, since they describe the lowest energy point of the PES, from which the rest of the PES can be explored. Low-energy structures are crucial for most chemical systems and can be created using reasonable random displacements, which introduce variance in the chemical environments without the need of prior knowledge of the system. A systematic mapping of all chemical environments is unfeasible due to the complexity of the system and the cost of the needed calculations. The data set needs to be restricted to the chemical environments available for the system given the nature of the system of interest. For example a structural configuration, which needs a reorganisation that

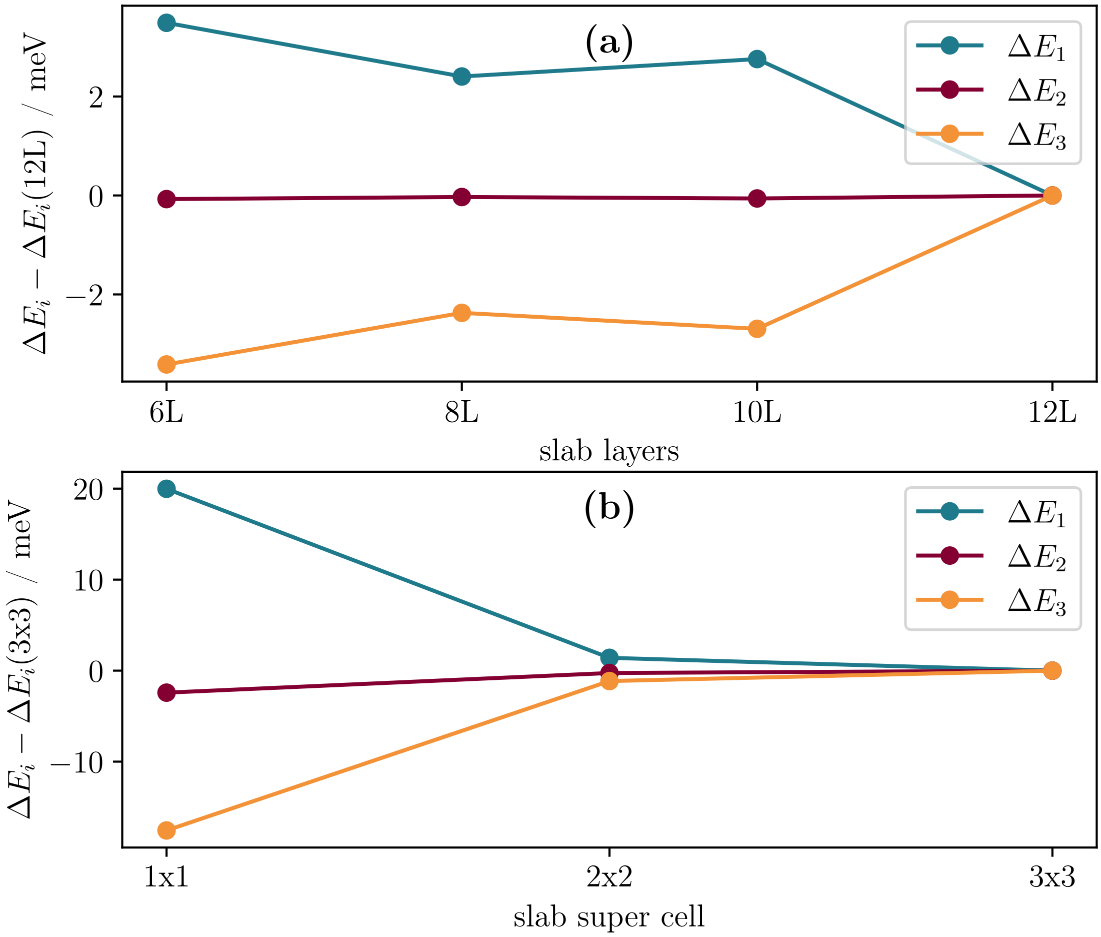


Fig. 3.4: The energy convergence for slab O-Al-Al layer count and super cell size. **(a)** Differences between three H atom positions above the α-Al_2O_3 slab with a different number of layers, from six layers as in the bulk unit cell up to double the amount. ΔE_1, ΔE_2 and ΔE_3 are the differences in energies between the positions of H atoms shown in figure 3.2. Energy convergence of less than 4 meV is reached with eight O-Al-Al layers, which also increase the distance between the two surfaces to more than twice the cutoff radius for the symmetry functions. **(b)** Differences between three H atom positions above the α-Al_2O_3 slab at different super cell sizes. Energy convergence of less than 1.5 meV is reached for the 2×2 super cell.

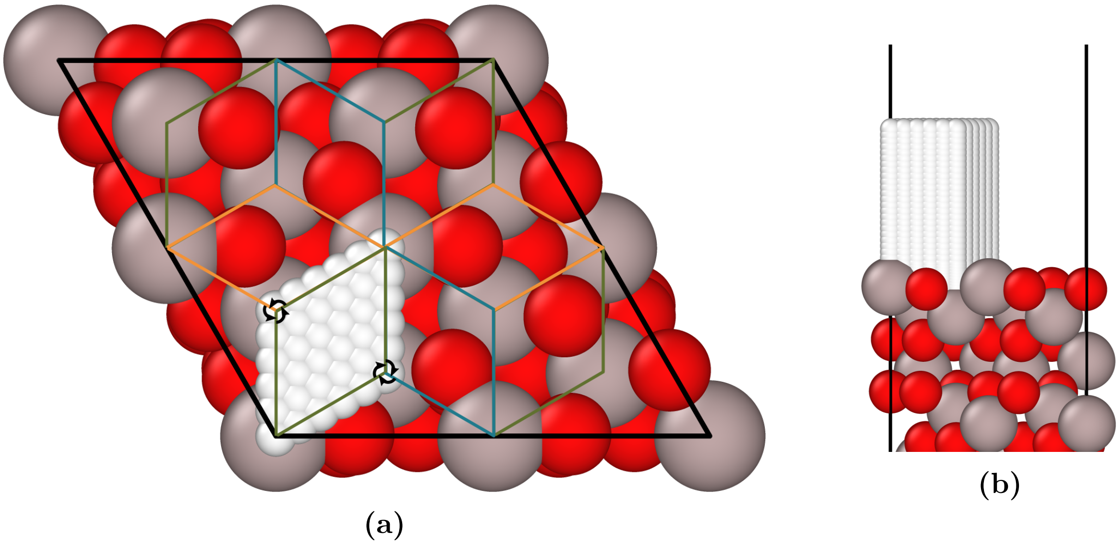
only happens at 1000 K, does not need to be included into a reference data set for a system used for simulations at 300 K. Small random displacements around the energetic ground state will enable the first HDNNP fit to run in the direction of thermal distortions. Reasonable displacements vary from system to system, but one should stay with small displacements. Especially when working with an incomplete data set, high-energy or force outliers can easily derail the fitting process. Here, displacements of 0.2 Å yielded reliable results, while higher displacements caused problems in the initial fitting process. Those random displacements were made for the optimized geometry structure, as well as for structures with scaled lattice vectors of up to ±2%. Structures with distances between the atoms of less than 1.2 Å were filtered out.

A similar process was applied to generate slab structures, without scaling the lattice constants and random displacements up to 0.2 Å. In order to add the first chemical environments for the H atom, a grid approach was applied. Initially, structures of the geometry optimized slab with H atoms in an equally spaced grid in the irreducible unit of the unit cell were calculated. Due to the symmetry of the surface the H atom only needs to be placed over one third of one unit cell to gain information about every possible chemical environment. The other two thirds can be obtained by three-fold rotation of the irreducible unit as shown in figure 3.5 alongside a representation of the grid. The spacing of the grid in all direction was initially set to 0.5 Å, from a distance of 0.6 Å above the surface, measured by the oxygen layer, up to 7.1 Å from the surface. Again, structures with atomic distances of below 0.75 Å were excluded. Later a closer spacing of 0.3 Å for layers between 0.6 Å and 1.7 Å was added to improve the description of the chemical environments close to the surface. The H atom was added on top and below the slab to maintain the inversion symmetry of the slab. To further increase the initial quality of the fit, H atoms were added above slab structures with randomly displaced atoms, similar to the slab structures without a H atom. Although the inversion symmetry was broken because of random displacements, H atoms were still added on both sides of the slab to gain more H atom chemical environments per structure.

After creating an initial fit using the structures mentioned in this section, additional structures were added using active learning which is described in more detail in section 3.3.1.

Initial Structure Generation for the First RPBE Data Set

The structures created for the first data set served as a catalyst to initiate the development of the second data set. Applying the same methodology used during the generation of the PBE data set, the geometrically optimized bulk and slab structures were calculated and scaling factors were determined to adjust the PBE lattice constants to match the RPBE lattice constants. Using these scaling factors, the lattice constants and atom positions of the structures in the PBE data set could be scaled to mimic RPBE structures. The scaled structures could then be calculated using the RPBE functional. This created a huge initial data set that resulted in an especially stable first fit. Scaling the structures using just the factor between the lattice constants would not have reproduced the subtle differences between both functionals, for example differences between the Al-O distances and the O-O distances, which would be scaled equally. But the scaled structures would still sample a reasonable structural space on the RPBE PES. This sampled space would be biased by the structures preferred by the PBE PES and extra care has to be taken to remove this bias by adding the missing


(a)

(b)

Fig. 3.5: **(a)** Top and **(b)** side view of the H atom grid used for the final RPBE data set. The red spheres are O atoms, the grey spheres are Al atoms and the white spheres are H atoms. In the top view the irreducible unit of the α-Al_2O_3 unit cell is framed in green. The black arrow markers show the three-fold rotation axis to rotate the irreducible unit to fill the whole unit cell. The orange and blue frames are the rotated irreducible units shown inside the 2×2 super cell of the slab structure.

structures needed to fully describe the RPBE PES. Especially any local minima in the RPBE PES, which are not in the PBE PES, need to be found and sampled. This structural search was simplified by already having a HDNNP fit which runs relatively stable MD simulations. Those missing structures then needed to be found applying active learning further described in section 3.3.1.

Initial Structure Generation for the Final RPBE Data Set

After detecting problems with the calculation of the electron spin, explained further in section 4.1.4, the general setup of the structures containing H atoms had to be changed. In summary, electron spin of the H atom on one side of the slab could be moved to the other side of the slab during the DFT SCF cycle, resulting in energy-structure relations which the HDNNP is not able to describe due to the cutoff of the symmetry functions. As a solution for this problem the lowest 4 O-Al-Al layer were frozen in geometry optimised places and only one H atom on the top side of the slab was added. This means that the top side of the slab, including the H atom, is always interacting with the same bottom side, which eliminates the need for the HDNNP to be able to see the other side of the slab. Since the bottom side is frozen in geometry optimized positions, which are energetically favored, the amount of interaction between both sides should be kept at a minimum. Analyzing the Hessian matrix of the slab, with the H atom in close proximity to the surface atoms, the residual atomic Hessian submatrix norm values $||\mathbf{h}_{\mathrm{HX}}||$, between the atoms in the first frozen layer and the H atom, were determined. The value is below $0.2\,\mathrm{eV}\,\mathrm{Å}^{-2}$ for every frozen atom, as shown in figure 3.6. As frozen layers are now required for all structures with a H atom, only structures lacking H atoms were eligible for transfer from the initial RPBE data set. A set of structures with the H atom on a grid was created using the procedure described in section 3.3.1. Additionally, the HDNNP fit, using the first RPBE data set, was used to generate the structures for the final data set. For this structure search, a farthest point sampling approach (FPS)[99] was applied. This was now easily applicable because there was already a working HDNNP fit and only one H atom per structure.

MD simulations under experimental conditions were run only adding one H atom and freezing the lowest four O-Al-Al layers using the HDNNP trained on the first RPBE data set. The symmetry function values of the H atom can be used to find the H atom chemical environments in the trajectory, which are the least similar to the chemical environments already included in the data set. To find the structures with the most different chemical environments for the H atom in the trajectory, the Euclidean distance of the difference between the symmetry function vector of the H atom in the new structure with every symmetry function vector of the H atoms in the data set was calculated. This results in a list of values corresponding to the similarity between the H atom chemical environment of the new structure with the chemical environments already included in the data set. The smallest of those values for every structure in the trajectory was then compared with each other. From this list, the structure with the highest value was added to the data set. Of all structures of the trajectory the H atom has the least similar chemical environment with all chemical environments in the data set. In the beginning this process was repeated for a trajectory, after adding the first structure to the data set and including the first added structure in the search for the least similar chemical environment. This process ensures that many different chemical environments for the H atom are added to the data set without the risk of adding the same chemical environment multiple times. Atoms close to the surface have large symmetry function values, since they see a lot of atoms in their cutoff range. Initially mainly those structures are added to the data set, since this also means their difference vector is usually longer than the vector of H atoms further from the surface, where their symmetry function values are closer to 0. This is a useful characteristic of this method because more chemical environments for the H atom close to the surface are needed because these structures have the highest energies and forces. Using the HDNNP fit based on the first RPBE data set, structures were generated until the fit using the new data set was able to run simulations without too many extrapolation warnings. Subsequently, the data set was improved using a combination of FPS and active learning, described in the next section.

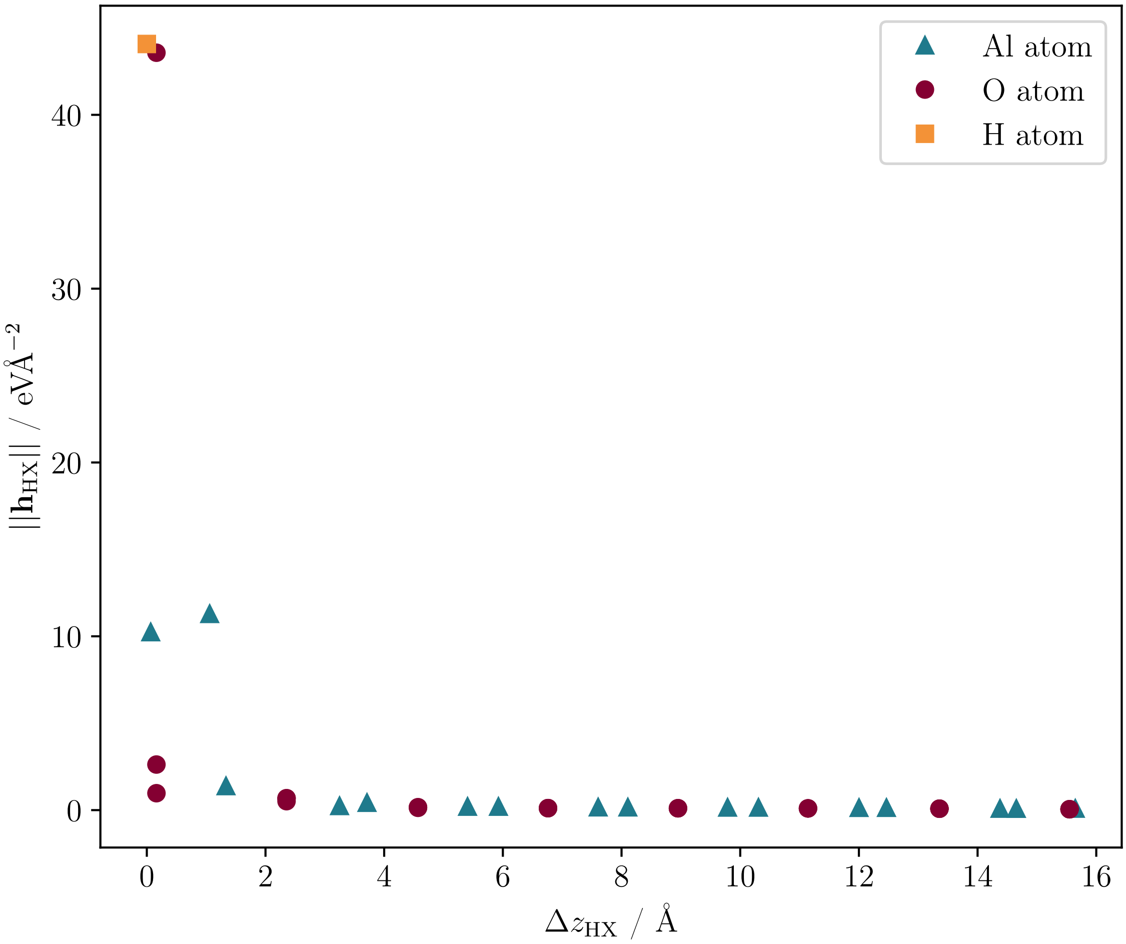


Fig. 3.6: Hessian analysis of the influence of atoms in the slab on the H atom in close proximity to the surface. The atomic Hessian submatrix norm values $||\mathbf{h}_{\mathrm{HX}}||$ for each atom is plotted vs the distances in z direction from the H atom. The frozen layers start at a distance of 8 Å.

Active Learning

In order to create a data set for machine learning potentials, there needs to be a way to generate structures which are missing to fully map the energetically relevant configuration space. This introduces the challenge that, especially when the goal is to create a reactive potential, not all possible configurations and pathways are already known. Often those potentials are created to explore pathways the system only rarely visits, which are impossible to explore using ab-inito MD, due to the high computational cost of long simulations. Here, active learning can be applied to use the potential to find missing structures and explore the configurational space guided by the forces and energies of the fitted PES. For active learning, HDNNP MD simulations are run in order to create new structures in the energy and temperature regions of interest. Then the structures are analysed and the structures which would expand the configurational

space of the data set are picked to add them to it. The data set can be extended to describe new regions of the PES by using the ACSF values to find structures with values not already included in the data set. Using the RuNNer program, those extrapolations are easily detected and the structures can be added to the data set.
Finding regions in the configuration space where the HDNNP does not have enough information to accurately interpolate is more difficult. Here, multiple HDNNP fits, using different architectures or splittings in training and testing set, can be used to calculate structures along the simulation to determine if their energy and force prediction diverge. Structures where two HDNNPs diverge in their prediction are not adequately covered by the data set and should be added to it. Using both approaches, the data set can be expanded and improved iteratively until the target accuracy and configurational coverage is reached. Especially early in the iterative process it should be checked that the structures found using active learning are not too similar. If simulations run into the same unexplored region of the PES and the predictions becomes too unreliable due to extrapolation, the extrapolated structures could be almost identical. It is recommended to limit the number of structures added for each iteration and increase the number of iterations. Increasing the data set size will also increase the fitting time, which then increases the time required per iteration. At this point the simulations should be stable enough to not run into the same local environment, where the amount of structures added per iteration can be increased. Running simulations at higher temperature can also speed up the mapping of relevant chemical environments after stable simulations are possible at the target temperature to generate missing structures that would take a long time to appear at the target temperature. Furthermore, structures should always be checked for unphysically short bond distances, since the HDNNP does not contain an energy penalty when atoms come very close. Setting reasonable distances for every element combination and discarding structures generated with too small distances can reduce the amount of unnecessary DFT calculations. This active learning method should be repeated until the HDNNP is at the desired accuracy and covers all configurations of interest.

Training of the High-Dimensional Neural Network Potential

In order to train a HDNNP using the reference data set the data set will be split into a training set and a test set. The fraction is usually 90% training and 10% testing set. The training set includes the structures the HDNNP fit will learn from, while the testing set is used to check the predictive power of the HDNNP and detect overfitting. Different settings for the atomic neural network architectures, the training algorithm, the weight initialization and the amount and weight of training on forces can be chosen to obtain an accurate HDNNP description of the PES. Depending on the system, some settings can have rather huge influences on the accuracy and training progress of the HDNNP and usually every hyper parameter should be tested at least once. Initially, the accuracy of the HDNNP fit can vary wildly, especially when fitting data with a large energy range or high forces. It is recommended to fit multiple HDNNPs per iteration to be more resilient against fits which are not converging, as well as to pick the best performing fits to use for active learning. Initially, the best criterion for fit quality is the RMSE of the energies, as well as the forces in the training and test set. The lower the RMSE, the better is the overall description of the training or test set. Usually the aim should be to have the smallest testing set RMSE possible. A small testing set RMSE indicates good transferability to unknown structures, which the test set is emulating. A big difference between training and test RMSE can be a sign of overfitting, where the interpolation quality of the fit is decreased, or due to an insufficient coverage of all chemical environments, where some chemical environments are only in the testing set. That indicates that the fit is not well trained for those structures and has high errors in the predictions. Usually those cases can be detected by plotting the error of the prediction vs the value of the energy or force. If there are structures with especially high error in the testing set but not in the training set it is a sign for an incomplete data set. Overfitting can usually be prevented by stopping the optimization process early, once the testing set RMSE stops decreasing. Plotting 1D energy curves along the PES can also help to find cases of overfitting by comparing a more densely populated HDNNP energy curve with the DFT reference calculations.

HDNNPs can also help improve the quality of the data set by helping to identify inconsistent electronic structure data. Finding single structures with small inconsistencies is nearly impossible because of the large number of structures in the data set and the thermal fluctuations of the structures already creating a wider energy range. But here simple grouping of structures can give enough additional information to detect those inconsistencies. For example the origin of the calculation can help finding calculations with different or wrong setups, or find a set of structures from MD simulations, all using a wrong starting structure. The easy accessibility makes plots of the prediction error against the binding energy or the structure number a great tool to check the data set. Another way to use the fitting power of the HDNNP to find inconsistencies in the data set is to use system parameters to split the data set and fit only a subsection of it. For example, one could compare the quality of fits including all structures with fits excluding especially skewed unit cells to find out if those calculations have convergence problems due to the skewed lattice angles. If the RMSE values of fits excluding those structures are far better than when including them, it can be very helpful to take a better look at those structures to determine if they are inconsistent with the rest of the calculations or just not sampled enough. In this project fits with all training data were compared to fits only including structures with a total spin of $S = \frac{1}{2}$ per H atom. Although spanning a similar energy range, the fits using only structures with a total spin of $S = \frac{1}{2}$ per H atom had roughly half the RMSE of the fits using all training data, as can be seen in figure 3.7. Further analysis showed that the data points with a spin of less than $S = \frac{1}{2}$ per H atom were not consistent with the rest of the structures, doubling the RMSE for fits including all data points.

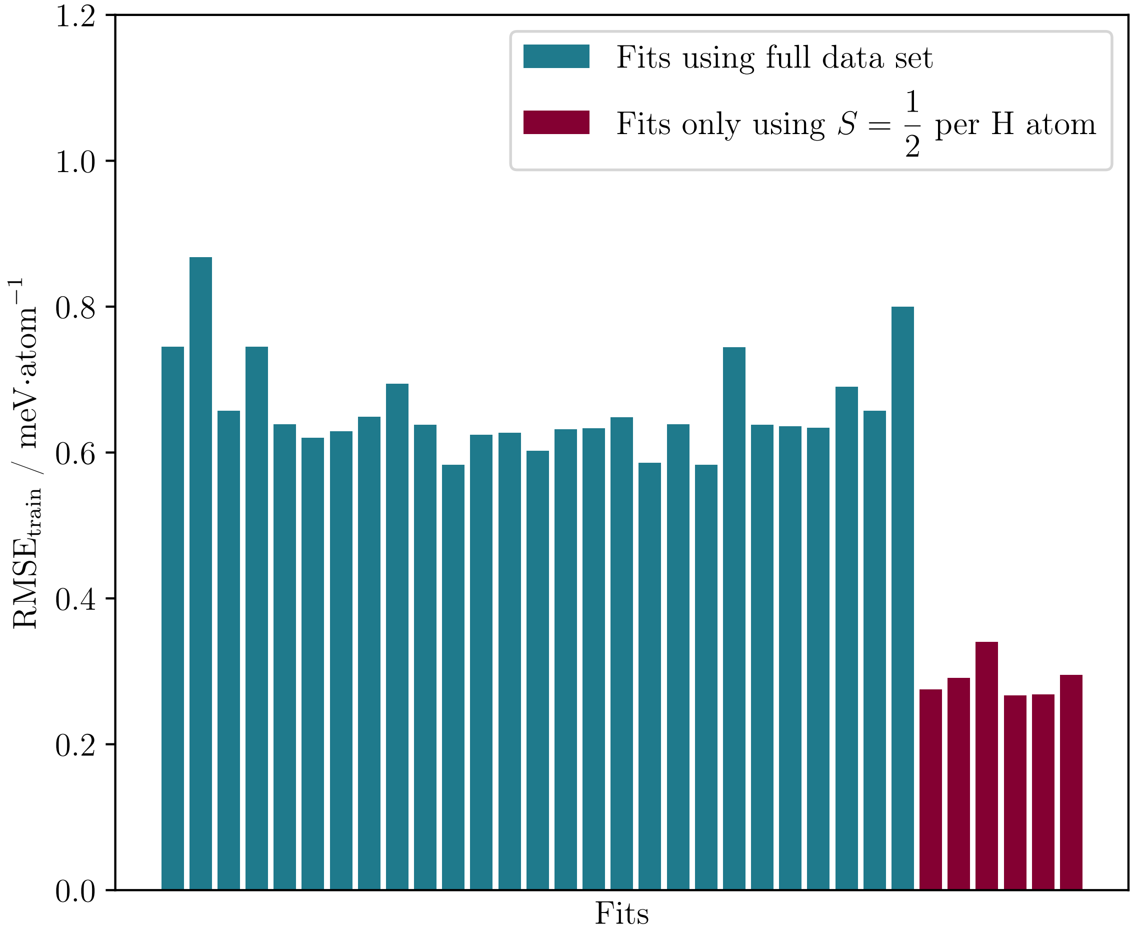


Fig. 3.7: Comparison of different train RMSE values for fits that include all structures in the PBE data set and fits that only use structures with a total spin of $S = \frac{1}{2}$ per H atom.

It is possible to establish upper limits for energy and force values to eliminate irrelevant configurations that are not relevant to the desired temperatures and pressures. Additionally, it is possible to set a different threshold for forces to exclude high forces from the fitting process, while keeping the structure and all other forces accessible for the fit. Those thresholds should be checked regularly to not exclude relevant data points during the sampling of all important regions of the PES.

3.3.2 RuNNer

The RuNNer code (version from August 22, 2019)[55,56,100] was used to construct the final RPBE HDNNP. A cutoff of $R_\mathrm{c} = 12\,a_0$ or 6.35 Å was applied to the ACSFs. The parameters of the radial ACSFs are compiled in table A.1. For the radial ACSFs all combinations for the parameters $\lambda = \{-1, 1\}$, $\zeta = 1, 2, 4, 16$, and $\eta = 0\ a_0^{-2}, 0.05\ a_0^{-2}$ were used. Some combinations of the radial symmetry functions are missing, which are listed in the appendix in table A.2. This results in 46 H ACSFs, 98 O ACSFs, and 94 Al ACSFs. Atomic neural networks consist of 2 hidden layers with 15 neurons each. The weights were initialized using the Nguyen-Widrow scheme. The HDNNP was trained on the total DFT energies minus the atomic energies, calculated using single atoms in vacuum. A full list of the RuNNer settings is provided in Tab. A.3.

4 Results and Discussion

4.1 Validation of the Final RPBE α-Al_2O_3 High Dimensional Neural Network Potential

4.1.1 Overview

With the aim of accurately describing the interaction between a H atom and the [0001] α-Al_2O_3 surface using a theoretical representation, the reference data and the HDNNP trained on the reference data need to be carefully validated. Only if the reference data set is well converged and without inconsistencies, the HDNNP can as accurate as possible. And only if the representation of data set by the HDNNP is as accurate as possible the resulting insights into the system can be trusted. To gain this trust the careful construction of the structural setup was described in section 3.3.1. Here the data set will be analyzed, using the HDNNP in order to find inconsistencies within it. The bulk α-Al_2O_3 system will be analyzed and the prediction of the HDNNP will be compared to the reference DFT results. The PES of the grid structures will be analyzed and the PBE and RPBE PES will be compared. The description of the bulk and surface α-Al_2O_3 systems is of great importance for the energy transfer from the H atom to the surface during the MD simulations. Comparing the phonon band structure calculated using the HDNNP with the band structure calculated with DFT will show how well the HDNNP is able to reproduce energy transfer mechanisms in the system. Using a structural setup to circumvent the H atom spin inconsistencies during the SCF cycle of the DFT calculation, the H atom interaction can be validated by comparing the DFT PES with the HDNNP PES in detail. Ab-initio MD simulations can also be recalculated using the HDNNP to ensure correct prediction during MD simulations using experimental conditions. Lastly the kinetic energy distributions calculated from simulations using the HDNNP can be compared to experimentally measured distributions at different incident conditions to not only validate the HDNNP but also investigate how well the functional is able to describe the system and understand where there are deviations from reality.

4.1.2 Reference Data Set

The final RPBE data set consist of 15812 structures, which include 808 bulk structures and 15004 slab structures from which 2300 do not contain a H atom. The final HDNNP has a RMSE for the testing data set of 0.746 meV atom^{-1} for the energy and 0.103 eVÅ^{-1} for the atomic force components. For the training set the RMSE values are 0.257 meV atom^{-1} for the energy and 0.111 eVÅ^{-1} for the force components. The largest energy error is 5.47 meV atom^{-1}.

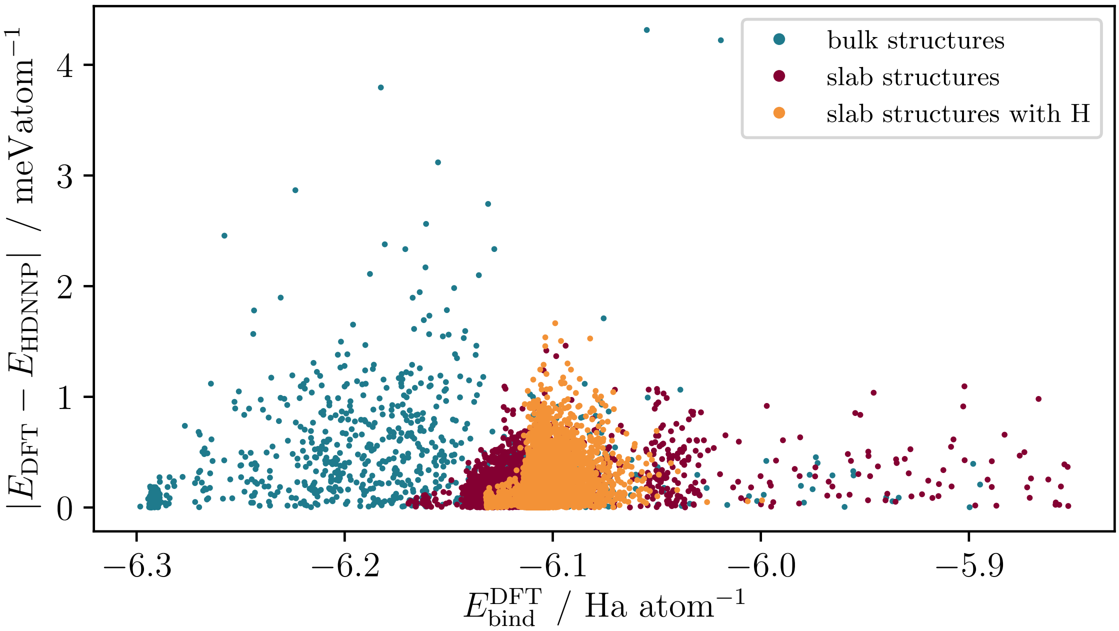


Fig. 4.1: Absolute prediction error $\Delta E = |E_{\mathrm{DFT}} - E_{\mathrm{HDNNP}}|$ of the training set plotted vs. the binding energy of the structure. The points are colored by type of structure, bulk, slab and slab with H atom. The lower the energy of each structure type the lower the errors, which tend to rise as the energy level increases.

The energy RMSE is well below the target error for machine learning potentials fitted to DFT reference calculations of 1.0 meV atom^{-1}. For all structures including H atoms the bottom half of the slab is frozen. These frozen atoms sum up to 41% of all atoms. Those atom share similar chemical environments between in all structures. The energy and force prediction for those atoms is very accurate, due to their abundance in the data set. Including such frozen atoms to the structures decreases their RMSE, without increasing the accuracy of the prediction for the moving region of the slab and the interaction with the H atom.

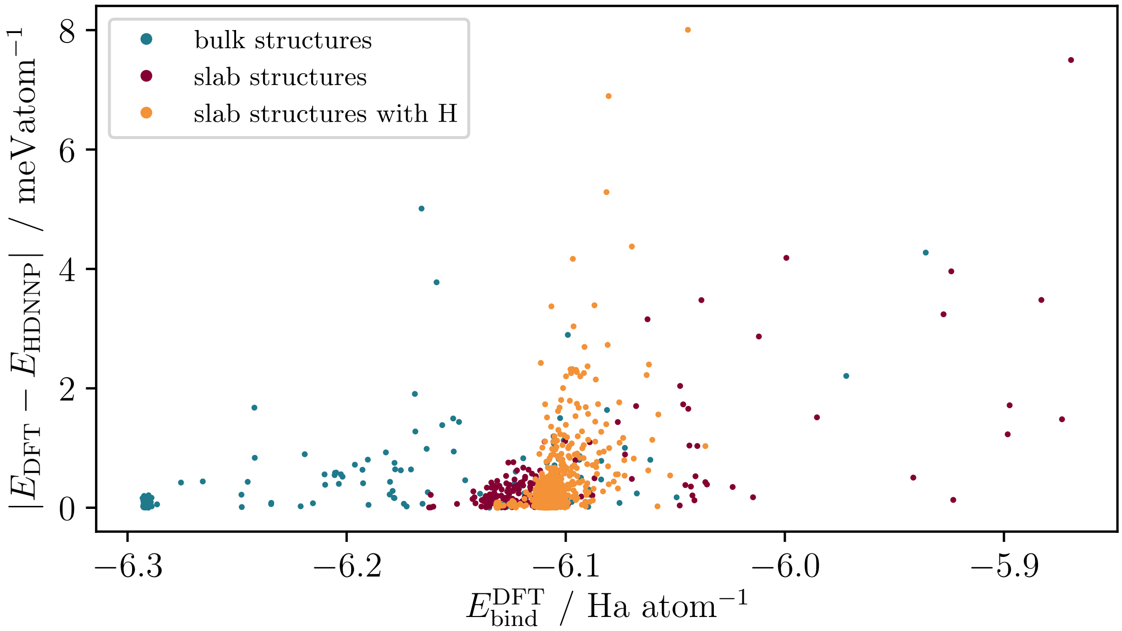


Fig. 4.2: Absolute prediction error $\Delta E = |E_{\mathrm{DFT}} - E_{\mathrm{HDNNP}}|$ of the testing set plotted vs. the binding energy of the structure. The points are colored by type of structure, bulk, slab and slab with H atom. The maximum prediction error is roughly twice as large as for the training set (figure 4.1), but most of the structures are described as well as in the training set.

4.1.3 Validation of the bulk System using Density Functional Theory

First the bulk α-Al_2O_3 structures will be analyzed. They are crucial for a accurate description of the target system, since the inside of the slab should be as close to the bulk system as possible. In order to teach the HDNNP the correct bulk energies and forces directly, the bulk structures are included in the training process. As a first step of validating the potential we can calculate the lattice constants and bulk modulus B_0 for bulk α-Al_2O_3, which are shown in table 4.1. The bulk modulus is calculated using the Birch-Murnaghan equation of state[101,102]

$$E = E_0 + \frac{B_0 V}{B'}\left[\frac{\frac{V_0}{V}^{B'}}{B'-1} + 1\right] - \frac{B_0 V_0}{B'-1}\ , \tag{4.1}$$

determining the equilibrium energy E_0 and volume V_0 as well as the bulk modulus as

$$B_0 = -V \left.\frac{\mathrm{d}P}{\mathrm{d}V}\right|_T\ . \tag{4.2}$$

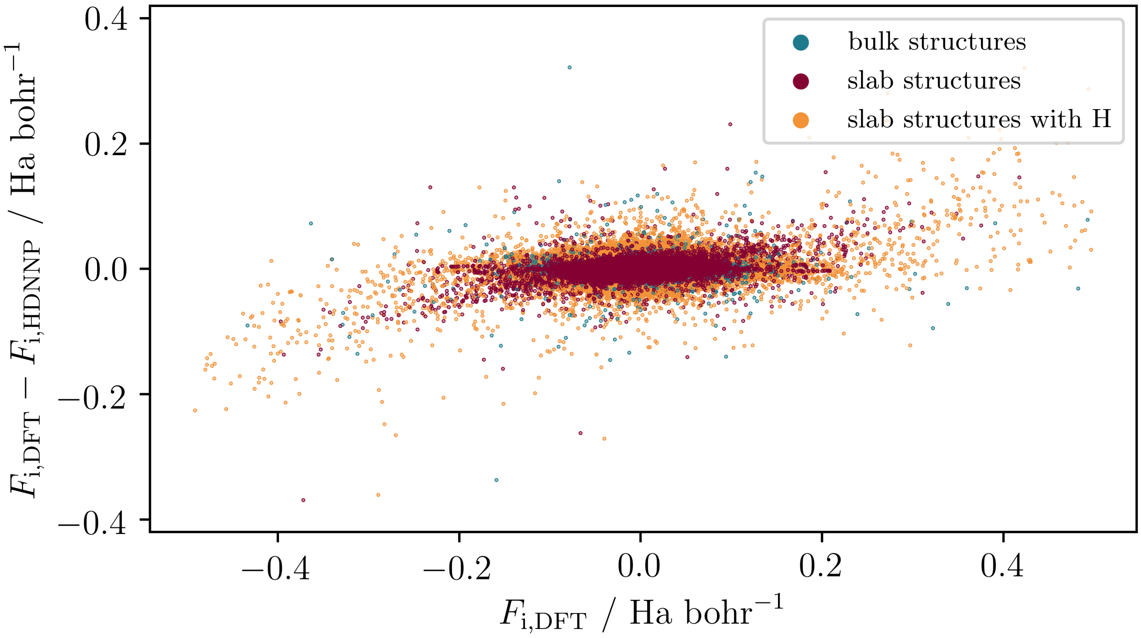


Fig. 4.3: Prediction error $\Delta F_i = F_{i,\mathrm{DFT}} - E_{i,\mathrm{HDNNP}}$ of the training set for each Cartesian force component i plotted vs. the DFT force component. The points are colored by type of structure, bulk, slab and slab with H atom.

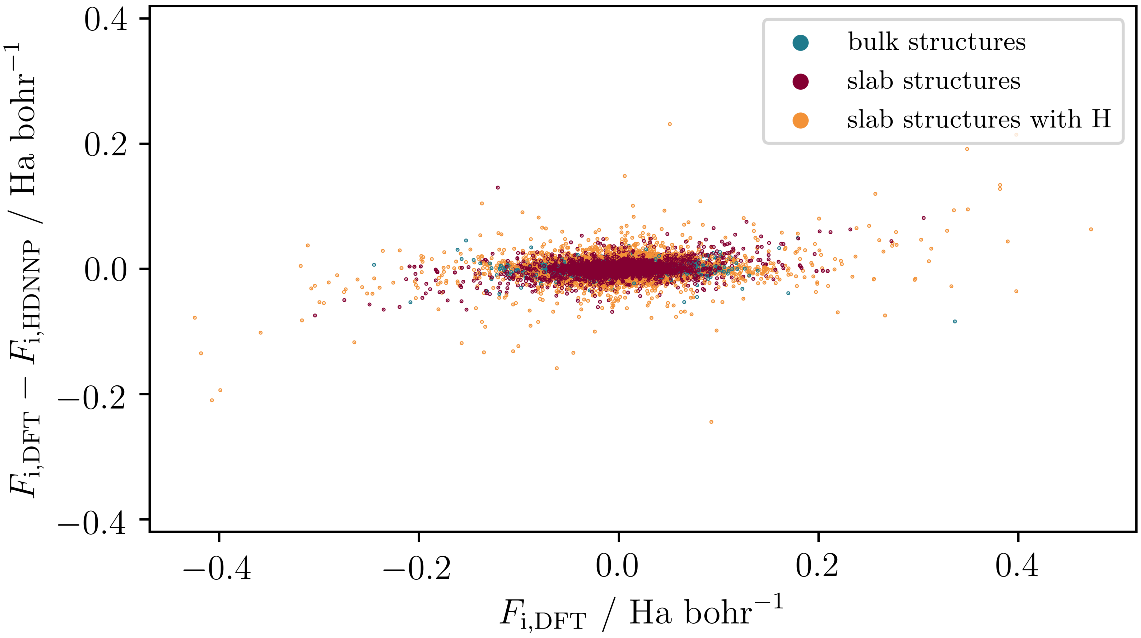


Fig. 4.4: Prediction error $\Delta F_i = F_{i,\mathrm{DFT}} - E_{i,\mathrm{HDNNP}}$ of the testing set for each Cartesian force component i plotted vs. the DFT force component. Force error distributions are very similar to the training set indicating good transferability.

The energy E and volume V are fitted for each $^{c}/_{a}$-ratio scaled from 0.97 to 1.03 in 0.01 steps. The HDNNP accurately reproduces the RPBE lattice constants for the

Tab. 4.1: The calculated and experimental lattice constants a and c of α-Al_2O_3[103] in Å and bulk modulus B_0 in GPa.

Compound	parameter	RPBE	HDNNP	Exp.
α-Al_2O_3	a	4.827	4.824	4.7554
	c	13.145	13.139	12.991
	B_0	222	216	253

bulk structure and predicts a slightly lower bulk modulus. The bulk lattice constants are used for the creation of the slab, here a good representation of the bulk properties is important for the description of interaction between the surface and the H atom. The deviation of the bulk modulus prediction is small and should result in a good representation of the compression characteristics of the system. The accurate prediction of the bulk lattice constants and bulk modulus lays the foundation for a accurate description of the slab system.

During the scattering process the H atom will loose kinetic energy, which is transferred to the surface in different channels. One of the channels for the H atom to transfer energy is phonon excitation. The accurate description of the phonons is important for the system in order to accurately describe the kinetic energy loss of the H atom during the scattering process and the energy distribution across the slab. This mechanism will greatly influence the final measured kinetic energy of the H atom, which makes an accurate description crucial for the comparison with the experiment. In Figure 4.5 the phonon band structures calculated using the phonopy program[96,97] is shown. The band structure calculated using DFT forces is compared to the band structure calculated using HDNNP forces, showing small differences between both spectra at low frequencies, which slightly increase at higher frequencies. The overall band structure is very similar and will result in very similar phonon excitation properties.

The HDNNP can accurately describe the bulk properties, which is shown by the small RMSE values, as well as the lattice constants, the bulk modulus and the phonon spectrum. This gives the HDNNP a ideal basis for the description of the α-Al_2O_3 slab, as well as the interaction with the H atom.

4.1.4 Investigation of the H Atom Spin

Overview

During the validation process of the HDNNP one of the system properties investigated thoroughly was the H atom spin. Calculations of structures including H atoms were initialized with a spin of $S = \frac{1}{2}$ on the H atom and analyzed using the Hirshfeld spin moments.[45] For H atoms far away from the surface the spin is located solely on the H atom, when the H atom is closer to the surface more and more spin density is

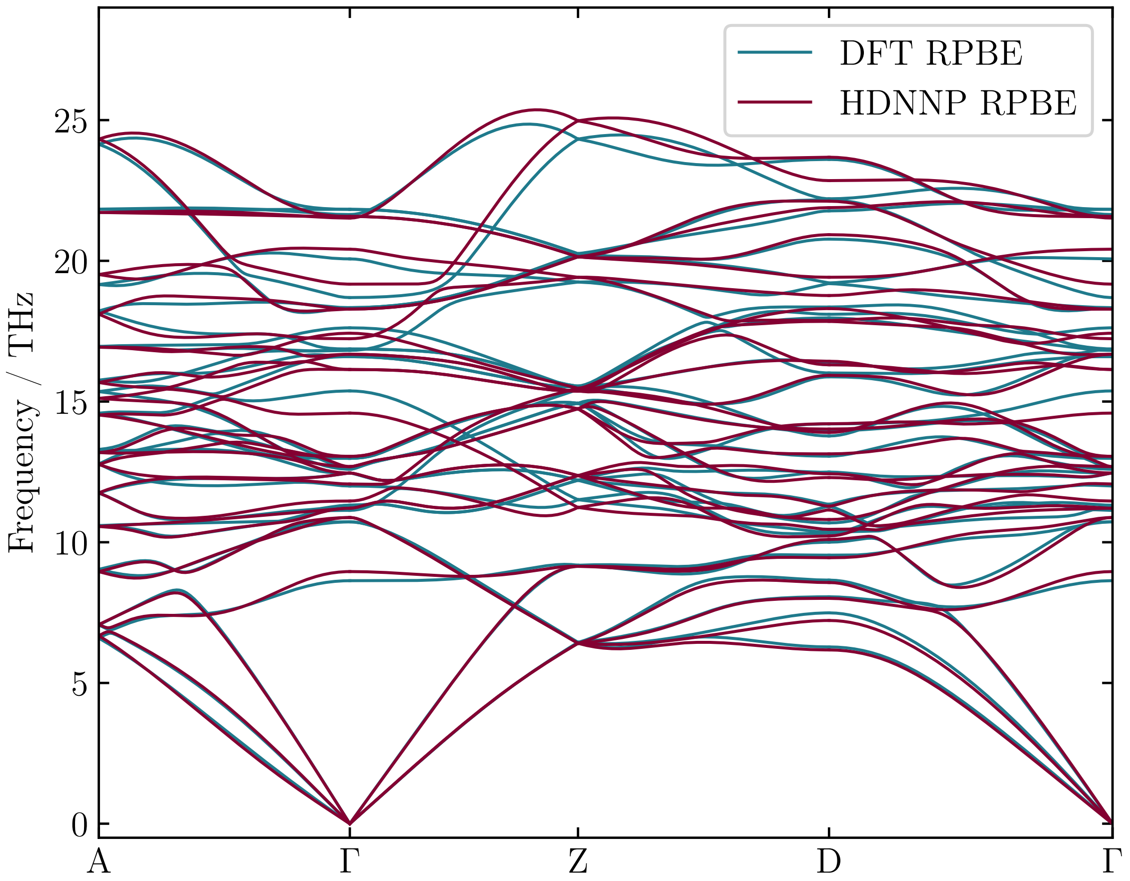


Fig. 4.5: Phonon band structure of bulk α-Al_2O_3 using a (2x2) primitive cell, created from the optimized 1x1 bulk unit cell. The spectrum was calculated using phonopy[96,97]

shared with the surface atoms, decreasing the spin density on the H atom. The general behavior is shown for the PBE data set in figure 4.6 and for the RPBE data set in figure 4.9. In order to analyze the total spin of the structure, the absolute total spin

$$S_{\mathrm{abs,tot}} = \sum_{i}^{N_{\mathrm{atom}}} |S_i| \tag{4.3}$$

is calculated as the sum of the absolute spin components of each atom $|S_i|$ in the structure. Based on the absolute total spin of the slab structures the DFT calculations can be classified into three different spin states. The calculation resulted either in an absolute total spin of $\frac{1}{2}$ per H atom, no spin, or a value between 0 and $\frac{1}{2}$ per H atom. During the SCF cycles of the calculation the electron density per spin state is constantly adjusted to find the minimum energy for the structure. During this adjustment the initial spin values are changed to lower the energy. If not specified directly, the total spin will also change during the calculation. For example both H atoms could start with a spin of $S = \frac{1}{2}$ and during the SCF cycle the spin of one H atom flips to $S = -\frac{1}{2}$, resulting in a total spin of the structure of 0, and an absolute total spin of $S_{\text{abs,tot}} = 1$. The majority of the computations resulted in a spin state with $S_{\text{abs,tot}} = \frac{1}{2}$ per H atom, whereas approximately 20% of them yielded a spin of 0 or a value in between these two. The following section will investigate this behavior to ensure data set consistency. The goal is to accurately predict the DFT results using the HDNNP.

Spins in the PBE Data Set

In order to investigate the interaction of the H atom with the surface the H atom spin was plotted as a function of the distance to the closest surface atom in figure 4.6. Additionally the data points are colored according to the absolute total spin of the structure. Most of the H atom spins follow a common pattern, at large distances from the surface the H atom carries the whole spin of the structure. When the H atom interacts with the surface the spin density is shared with the surface, reducing the spin of the H atom. In some cases the absolute total spin is also reduced, resulting in either some residual spin density or 0 spin density. A considerable amount of structures with H atoms closer to the surface end up with less than $S = \frac{1}{2}$ per H atom. Since the total spin of the system during a DFT calculation is not constant, it is possible that those calculations are in the correct DFT spin state. That would mean spin state is consistent in energy and forces with the rest of the calculations with a spin state of $S = \frac{1}{2}$ per H atom.

The number of structures that are not in the $S = \frac{1}{2}$ per H atom state is too vast to examine each one meticulously, and a preliminary random structural investigation was inconclusive. however, using the ACSFs the structures can be analyzed using the same method the HDNNP is using to represent the structure. There should be a correlation between the spin of a H atom and its symmetry function values, because both depend directly on the chemical environment. Plotting the difference between spin of two H atoms ΔS_{H} vs. the Euclidean distance between the symmetry function vectors

$$\Delta G_{\text{H}} = |\mathbf{G}_{\text{H}_1} - \mathbf{G}_{\text{H}_2}| \tag{4.4}$$

of the H atoms H_1 and H_2 should reveal this correlation. Figure 4.7 shows these plots in panels **(a)**, **(b)** and **(c)**. When two H atoms from structures with different $S_{\text{abs,tot}}$ are compared, they will be shown in panel **(b)** if one of them comes from a structure with $S_{\text{abs,tot}} = 0$ and none of them has a $S_{\text{abs,tot}}$ in between 0 and $\frac{1}{2}$ per H atom and, in panel **(c)** if either one of them has a $S_{\text{abs,tot}}$ in between 0 and $\frac{1}{2}$ per H atom. For structures with $S_{\text{abs,tot}} = \frac{1}{2}$ per H atom there is a clear correlation between the H atom spin and its symmetry function vector as seen in panel **(a)** of figure 4.7. Here, a different spin value always indicates a different chemical environment, described by the symmetry function vector. This is not the case for structures with an absolute total spin of less than $S = \frac{1}{2}$

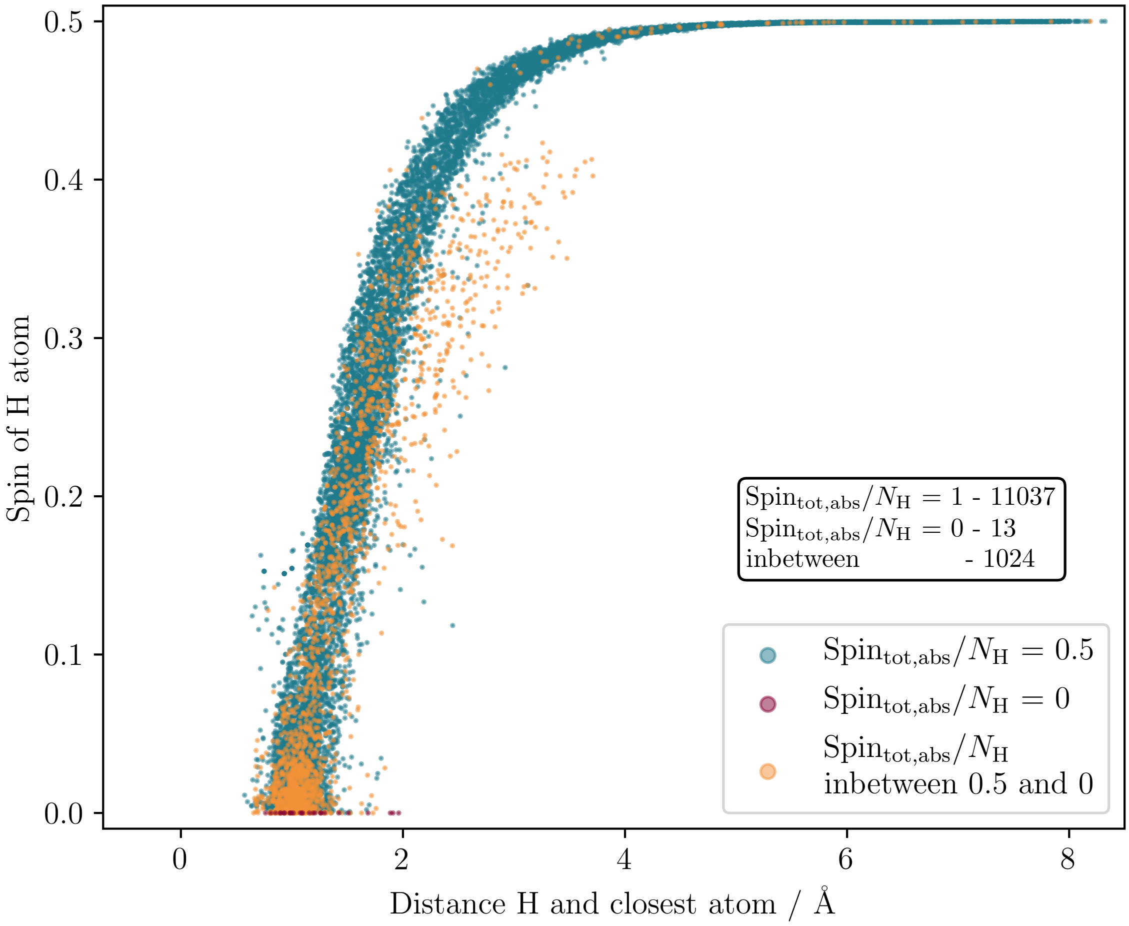
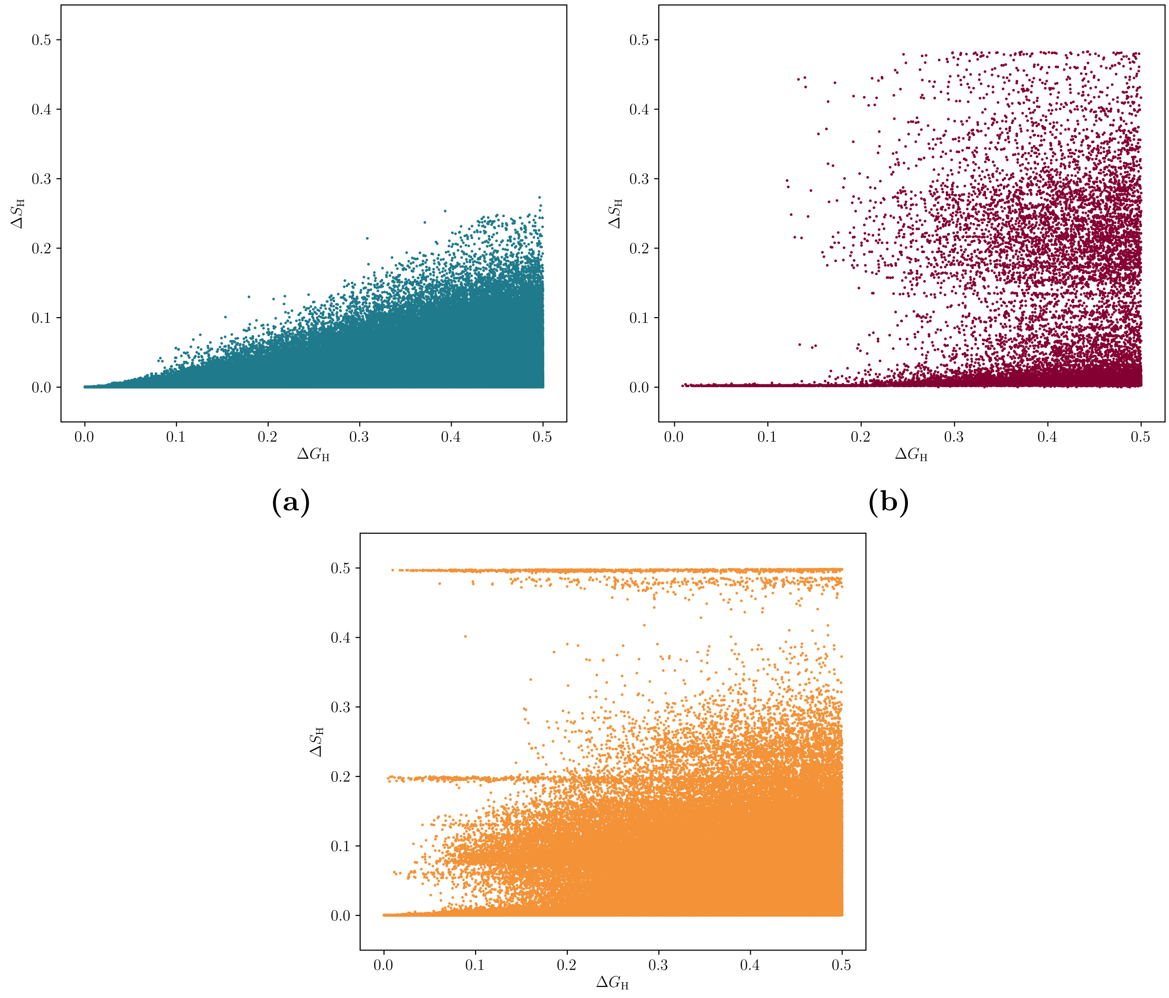


Fig. 4.6: H atom spin plotted over the distance between the H atom and the closest atom. The points are colored according to the absolute total spin of the structure, if the structure has a spin larger than 0.475 per H atom it is blue, less than 0.025 per H atom it is red and in between those to values it is yellow.

per H atom, as seen in panels **(b)** and **(c)** of figure 4.7. Here, even for very small ΔG_{H} the maximum difference in spin of $\Delta S = \frac{1}{2}$ is reached. Comparing all panels of figure 4.7 shows that the HDNNP is able to differentiate between the different spin values for the structures with $S_{\mathrm{abs,tot}} = \frac{1}{2}$ per H atom, while there are structures without a clear correlation between the chemical environment described by the ACSFs and the H atom spin with $S_{\mathrm{abs,tot}} < \frac{1}{2}$ per H atom. This lack of correlation presents a significant problem. If the HDNNP cannot distinguish between structures that possess varying energy-relevant characteristics, it will inevitably fail to make accurate predictions for those structures. This remains true even if such structures are consistent with the rest of the calculations. Ultimately, the HDNNP will prove incapable of reproducing the PES.

In order to test if those structures are problematic for the HDNNP, two sets of HDNNP



(a) **(b)**

(c)

Fig. 4.7: The difference in H spin ΔS_{H} between two H atoms plotted the Euclidean distance the symmetry function vectors of both ΔG_{H} shown for combination of H atoms from structures with a absolute total spin of $S_{\mathrm{abs,tot}} = \frac{1}{2}$ per H atom in **(a)**, combinations including H atoms from structures with a spin of $S_{\mathrm{abs,tot}} = 0$ and $S_{\mathrm{abs,tot}} = \frac{1}{2}$ per H atom in **(b)** and combinations including H atoms from structures with a spin in between $S_{\mathrm{abs,tot}} = 0$ and $S_{\mathrm{abs,tot}} = \frac{1}{2}$ per H atom in **(c)** for all H atoms in the PBE data set up to $|\Delta G_{\mathrm{H}}| = 0.5$.

fits were generated. The first set is fitted to the whole data set, including all structures, while the second set only considers structures with a spin of $S_{\mathrm{abs,tot}} = \frac{1}{2}$ per H atom, which also includes spin-free bulk and slab structures without H atoms. Comparing the generated fits will show the impact of the structures with $S_{\mathrm{abs,tot}} < \frac{1}{2}$ per H atom on the accuracy of the HDNNP. To compare the fits, their respective RMSE of the training set, which serves as an overall measure of accuracy for all structures, and the prediction error for an optimized slab structure with a H atom located beyond the cutoff radius of the surface are evaluated. The latter metric was chosen, because 1D energy plots of the adsorption of the H atom onto the optimized surface were investigated as a measure for fit accuracy, and an energy offset was observed at distances larger than the cutoff distance. This energy difference is indicative of a problem with the data set, since the HDNNP should, in principle, be able to learn the correct energy for a H atom and the optimized slab structure. The fits share the same neural network architecture and general settings and only differ in their random seed and the allowed energy range of the structures.

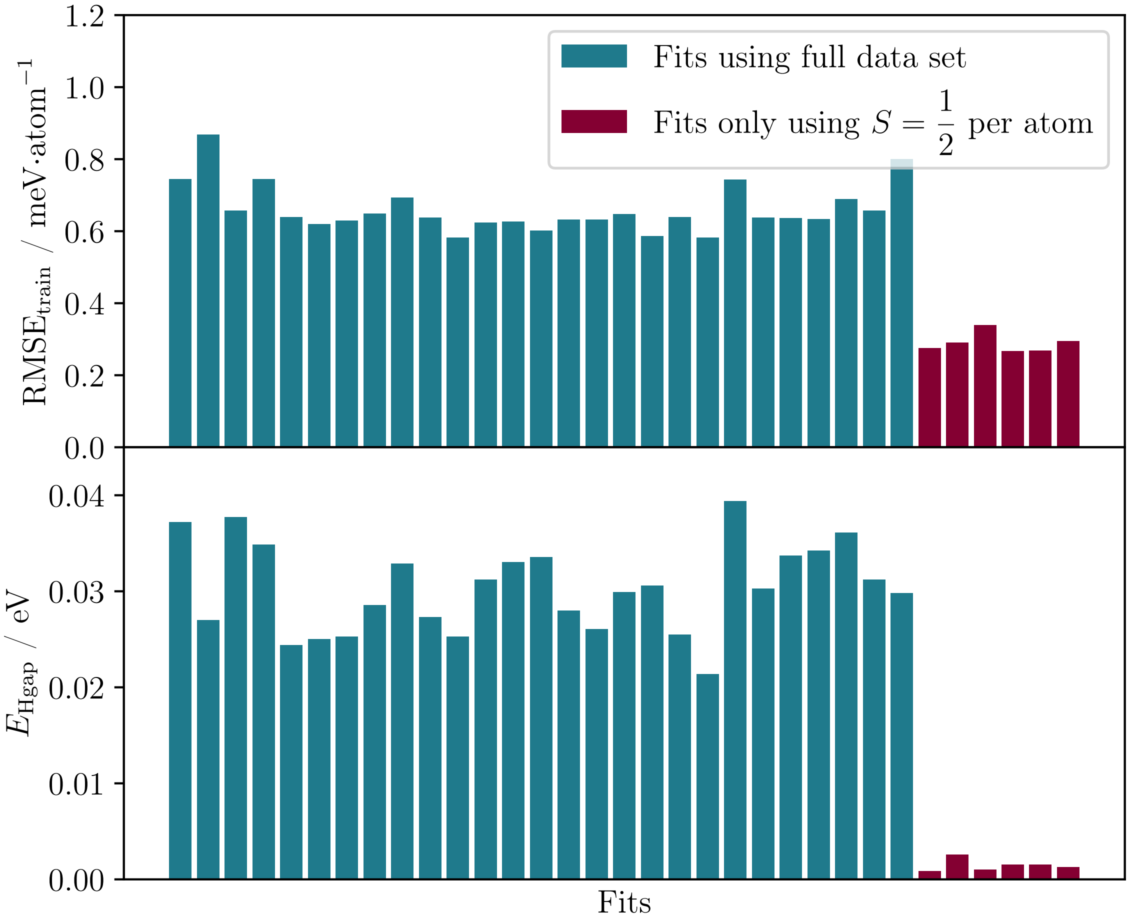


Fig. 4.8: Comparison of different training set RMSEs and the prediction error for an optimized slab structure with a H atom located beyond the cutoff radius of the surface E_{Hgap} for two sets of fits. The first set includes all structures in the PBE data set, whereas the second set only uses structures with an absolute total spin of $S_{\mathrm{abs,tot}} = \frac{1}{2}$ per H atom.

The results of both sets of fits are shown in figure 4.8. As can be seen, excluding the structures with a total spin of less than $S_{\mathrm{abs,tot}} = \frac{1}{2}$ per H atom reduces the mean RMSE of the fits from around 0.6 meV atom^{-1} to around 0.3 meV atom^{-1}, while also reducing the energy difference for structures of the optimized surface with H atoms at large distances from around 0.03 eV to less than 0.005 eV. The decrease in RMSE and removal of the energy difference at large H atom distances provides evidence that there is indeed a consistency problem between the structures with a total spin of $S_{\text{abs,tot}} = \frac{1}{2}$ per H atom and structures with less total spin, but the extent of the problem is still unknown. In particular, it is plausible that either all or only a specific subset of those structures are problematic. To efficiently train accurate models, there needs to be a way to ensure consistent data for the already calculated structures, as well as for all future structures without hindering the active learning procedure. Ensuring consistent data is crucial for constructing a machine learning potential and identifying the inconsistent data is a difficult and time consuming task. The chance of adding inconsistent data must be minimized to not cause problems in the future. First, the DFT calculations have to be checked for correct convergence. If the calculations are already correctly converged to lowest energy for each structure, the setup of the HDNNP or the structures needs to be changed. Since the calculations with an absolute total spin of $S_{\text{abs,tot}} = \frac{1}{2}$ per H atom are consistent, fixing the total spin of the system during the DFT calculation could lead to consistent data.

Fixed Total Spin Calculations

The HDNNP fitting power has revealed that structures with a spin value of less than $S_{\text{abs,tot}} = \frac{1}{2}$ per H atom are inconsistent with the rest of the structures. This inconsistency could stem from one of two issues. It is possible that the DFT calculation's SCF cycle did not converge to the same PES for every calculation, or that the HDNNP is incapable of distinguishing between structures with varying spin values, or perhaps both. Using the symmetry function values for the H atom, it was already determined that the HDNNP is not able to differentiate the structures, but it is not known if that is due to incorrect DFT calculations. With the aim to get a better understanding of the DFT calculation, structures were recalculated fixing the total spin to be $S_{\text{tot}} = \frac{1}{2}$ per H atom. If the resulting energies and spins are consistent with the earlier calculations which had an absolute total spin of $S_{\text{abs,tot}} = \frac{1}{2}$ per H atom, and create spin states which are discernible by the HDNNP, the problem would be solved. Depending on the structure, fixing the spin had different results. Some of the structures converged to reasonable spin states which where very similar to the structures with an absolute total spin of $S_{\text{abs,tot}} = \frac{1}{2}$ per H atom, but had a lower energy than the original calculation, whereas other calculations had a higher energy than before. Some calculations exhibited spin states, where the sum of all the atomic spin contributions of one side of the slab stayed below $S = \frac{1}{2}$ while the other side summed up to more than $\frac{1}{2}$. Further investigation of the latter case shows, that during the calculation the spin density is transferred from one side of the slab to the other, if that spin state is energetically favored. This is problematic in two different aspects. First this effect is invisible to the HDNNP. In order to correctly describe this spin shift, the atoms where the spin was shifted to would need to have the opposite side of the slab inside their cutoff radius. This would be against the idea behind this setup, which is to have two separate surfaces per structure, which are not influencing each other. This also is important for the second aspect of why this behavior is problematic. In a real system, the electron density cannot be shifted to the other side of the slab, because the distance between the two sides is too large. However, in some cases, this can artificially lower the energy of a structure that would otherwise be energetically unfavorable by moving the spin density to the other side of the slab, where it is less unfavorable. When using the HDNNP to describe such a system accurately without increasing the cutoff radius, only one hydrogen atom can interact with the surface, and everything invisible to this interacting H atom must be identical for all structures. If the side of the slab with the H atom interacts with the an identical opposite side of the slab for all structures, the effect of the other side of the slab on the hydrogen atom remains constant and does not require fitting by the HDNNP. To ensure this constant influence, the atoms on the side without the H atom need to be constrained in geometrically optimized positions. An energetically favorable structure will also limit the amount of spin density that can be shifted to this side. The exact changes to the structural setup are described in section 3.3.1.

Spins in the Final RPBE Data Set

Applying these changes to the structures and generating a new data set solved the apparent problem of inconsistencies in the data due to the spins. While the number of structures generated with $S_{\mathrm{abs,tot}} = \frac{1}{2}$ per H atom did not change, the HDNNP is now able to differentiate the different spin states. In figure 4.9 the H atom spin is plotted vs. the distance between the H atom and the closest atom on the surface. Previously, many structures with H atoms exhibiting spins $S_{\mathrm{abs,tot}} < \frac{1}{2}$ at higher distances and with relatively large H spins were present. Now, in contrast they are limited to structures where the H atom is in close proximity to the surface atoms and has lost most of its initial spin. The number of outliers also drastically decreased.

For this new data set, we perform the same test as for the PBE data set, to see whether the HDNNP can differentiate the spin states by plotting the difference in spin ΔS_{H} vs. the Euclidean distance of the symmetry function vectors ΔG_{H}. For all pairs of H atoms, symmetry function vector changes and H atom spin changes are correlated. In figure 4.10 panel **(b)** the number of points drastically decreased compared to the PBE data set. This is because the structures with an absolute total spin of $S = 0$ are in very different chemical environments compared to structures with different absolute total spins and the plot only includes pairs with $\Delta G \leq 0.5$. The tests show that the data set is more consistent with itself, because here different H atom spins are correlated with different chemical environments.

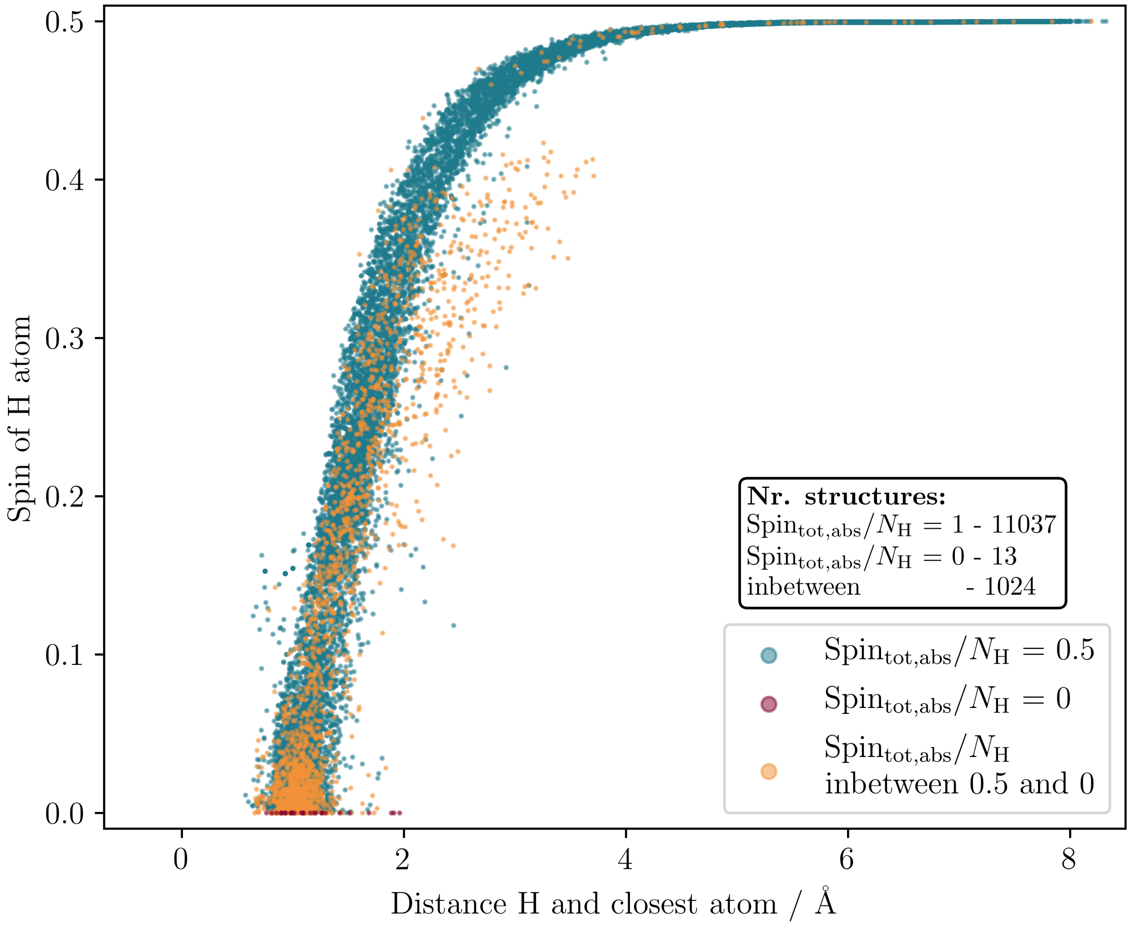
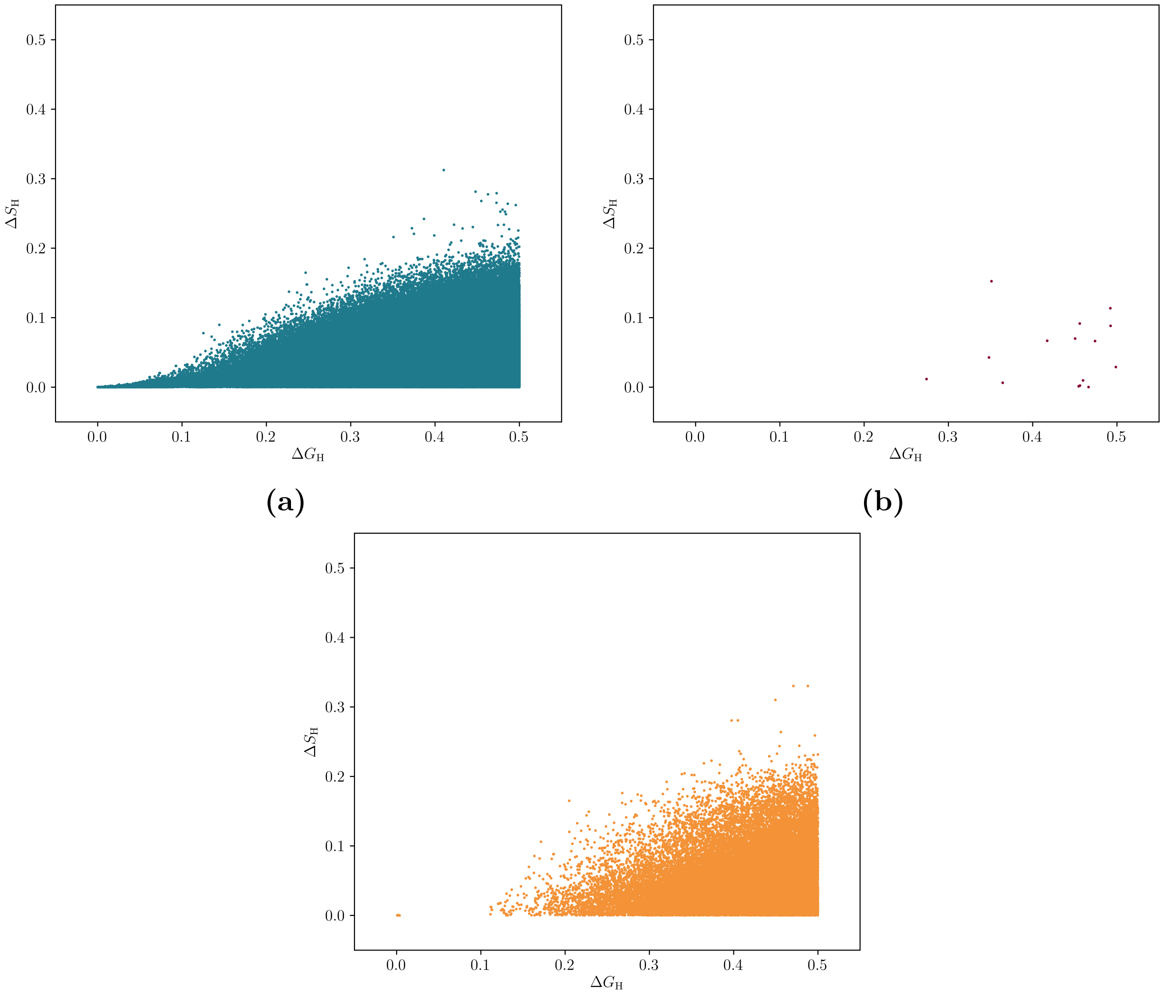


Fig. 4.9: H atom spin plotted over the distance of the H atom and the closest surface atom. The points are colored according to the absolute total spin of the structure, if the structure has a spin larger than 0.475 per H atom it is blue, less than 0.025 per H atom it is red and it is yellow otherwise.

As a final consistency test, again, fits including all structures and fits including only structures with $S_{\text{abs,tot}} = \frac{1}{2}$ per H atom, including bulk and slab structures, were calculated. The results of this test are shown in figure 4.11. The accuracy of the HDNNPs is the same for both data sets.

The quality of the data set is fundamental for the accuracy of the HDNNP. During the validation of the accuracy of the HDNNP it is important to investigate the potential in multiple different ways and not only focus on the target parameters. Since the H atom spin is not the target parameter, it was not the training objective to accurately predict the spin of the H atom, but analyzing how the spin changes with the distance of the H atom to the surface made it possible to detect problems with the DFT calculations and with the ability of the HDNNP to differentiate between different spin states.



(a) **(b)**

(c)

Fig. 4.10: Plot of the difference in H spin ΔS_H between 2 H atoms over the Euclidean distance of the symmetry function vectors of both ΔG_H. Pairs of H atoms from structures with a absolute total spin of $S_\mathrm{abs,tot} = \frac{1}{2}$ per H atom in **(a)**, pairs including H atoms from structures with a spin of $S_\mathrm{abs,tot} = 0$ and $S_\mathrm{abs,tot} = \frac{1}{2}$ per H atom in **(b)** and pairs including H atoms from structures with a spin in between $S_\mathrm{abs,tot} = 0$ and $S_\mathrm{abs,tot} = \frac{1}{2}$ per H atom in **(c)**, for all H atoms in the PBE data set up to $|\Delta G_\mathrm{H}| = 0.5$.

4.1.5 Comparing PBE and RPBE H Atom Interactions

The DFT functional determines the results of every calculation used in the data set for the HDNNP. Therefore, the choice of functional is therefore critical for an accurate description of the system. Within GGA functionals, there are two options, the PBE and RPBE functional. The first data set was created using the PBE functional. This choice was based on the fact that the lattice constant and bulk modulus of the bulk system were closer to the experimental values.[75] As previously stated, the bulk system forms the foundation for the slab system, with the slab being derived from a bulk structure.

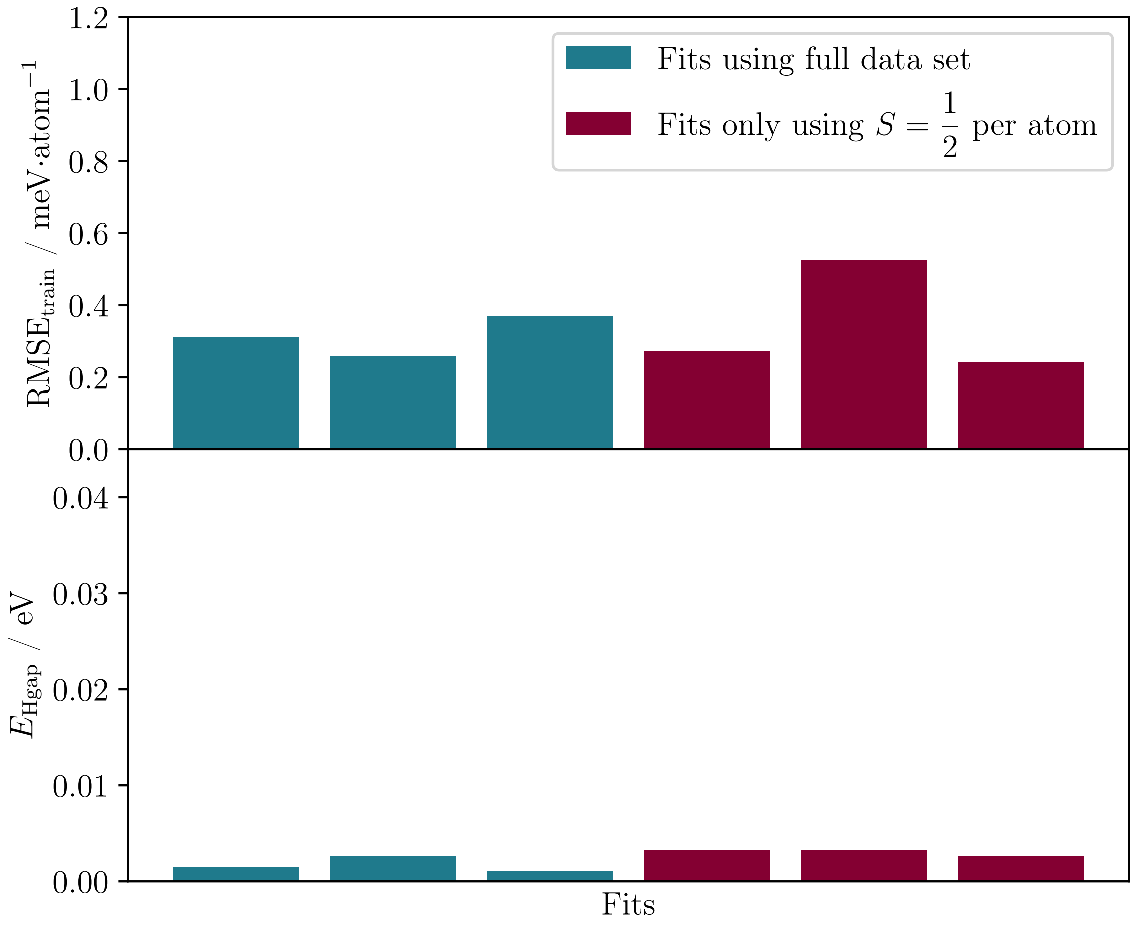


Fig. 4.11: Comparison of different training set RMSE values and the prediction error for an optimized slab structure with a H atom located beyond the cutoff radius of the surface E_{Hgap} for fits that include all structures in the RPBE data set and fits that only use structures with a total spin of $S = \frac{1}{2}$ per H atom.

The functional was selected based on the bulk properties. Once the HDNNP was created for the PBE data set, experimental kinetic energy loss spectra were compared to theoretical spectra, revealing a systematically higher kinetic energy loss predicted by theoretical calculations. This discovery prompted an investigation of the interaction between the H atom and the surface using both functionals to determine whether the RPBE functional would lead to a lower kinetic energy loss. Based on the experimental analysis of the system, the kinetic energy loss of the H atom is mainly driven by phonon excitation and can be described using the binary collision model.[104,105] The final kinetic energy of the H atom then can be approximated by the equation

$$E_{\mathrm{kin,s}} = E_{\mathrm{kin,i}} \cdot \frac{1}{\left(1 + \frac{m_{\mathrm{H}}}{M_{\mathrm{sur}}}\right)^2} \cdot \left[\sqrt{1 - \left(\frac{m_{\mathrm{H}}}{M_{\mathrm{sur}}}\right)^2 \cdot \sin^2(\theta)} - \frac{m_{\mathrm{H}}}{M_{\mathrm{sur}}} \cdot \cos(\theta)\right]^2 , \qquad (4.5)$$

where M_{sur} is the effective mass of the surface, m_{H} is the mass of the H atom and $\theta = \theta_{\mathrm{i}} + \theta_{\mathrm{s}}$. Following this equation, there are two factors influencing the kinetic energy loss, the effective mass of the surface M_{sur} and the initial kinetic energy $E_{\mathrm{kin,i}}$. Although the initial kinetic energy of the calculated trajectories matches the experimental value, the H atom can accelerate towards the surface, resulting in increased kinetic energy during the scattering event. This increase in kinetic energy leads to a larger kinetic energy loss compared to cases in which the H atom is not accelerated. After the scattering incident, the H atom loses kinetic energy as it moves away from the surface and exits the attractive potential of the surface. A more attractive potential thus increases the kinetic energy at the scattering incident, which then increases the lost kinetic energy. For both functionals, using the grid approach introduced in section 3.3.1, the PES for the H atom on top of the surface can be plotted. A comparison of different 2D cuts through the PES shows similarities and differences of the system with the two functionals.

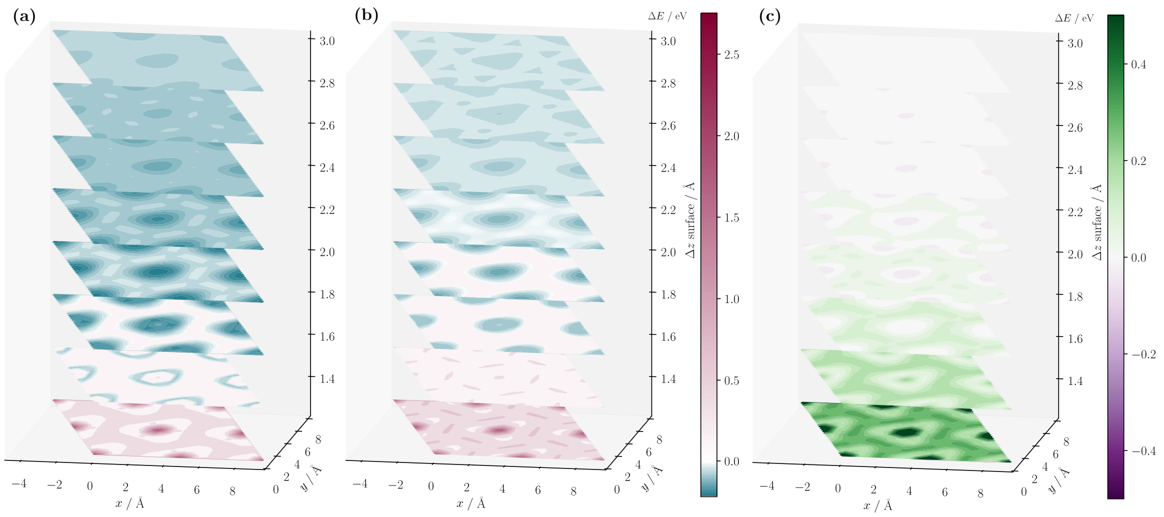


Fig. 4.12: 2D energy cuts through the PES of the H atom above the equilibrated surface. **(a)** depicts the PBE PES, **(b)** shows the RPBE PES and **(c)** shows the difference between both methods, $\Delta E = E_{\mathrm{RPBE,form}} - E_{\mathrm{PBE,form}}$. The energy difference ΔE for the PBE and RPBE energies is the difference between the optimized slab structure with the H atom at infinite distances compared to the H atom at the distance z from the O layer of the slab. The data points are from the grid approach described in the construction of the data set in section 3.3.1, and the points are interpolated linearly on a trigonal grid.

When comparing the adsorption energy of the H atom using the PBE and RPBE functionals (see figure 4.12), it becomes apparent that the RPBE surface is only half as attractive as the PBE surface. Specifically, the RPBE surface has an adsorption energy of 0.1 eV, while the PBE surface has an adsorption energy of 0.2 eV. If the H atom in the RPBE functional is more weakly bound to the surface, the kinetic energy

loss should be smaller.
However, also the other part of equation 4.5 also needs to be considered. The effective mass of the surface is determined by how strongly bound the surface atoms are and how well they can be deformed. The bulk modulus measures how well the bulk material deforms under pressure. Comparing the bulk modulus calculated with the RPBE functional, $B_0 = 222\,\mathrm{GPa}$, with the PBE functional, $B_0 = 228\,\mathrm{GPa}$, shows that the bulk modulus for the RPBE functional is not significantly lower than for the PBE functional. But the relative difference is much smaller compared to the difference in adsorption energy. While the lattice constants and bulk modulus of the PBE functional are slightly closer to the experimental values, the H atom adsorption energy of the RPBE functional is only half of the PBE functional. Considering these factors contributing towards the H atom energy loss discussed in this section, the clear choice of functional is the RPBE functional. The rest of the thesis will fully focus on the RPBE results.

4.1.6 Validation of the Slab System using Density Functional Theory

Overview

Ensuring consistent reference data is one step of validating the accuracy of the HDNNP, as the HDNNP can only predict accurately if the the data is consistent. If the reference data is consistent the next step is to validate that the potential is accurate in the relevant regions of the PES. Therefore, it has to be investigated, if the structures in the data set cover the relevant regions and allow for an accurate prediction of all relevant chemical environments. A general view of the PES can be gained by comparing the prediction of the HDNNP with the DFT energies for the grid structures. This will give a basis for the accuracy of the predictions over a larger structural space at nearly no extra cost. It can lay the base line of the expected accuracy, since it shows the PES for the lowest energy structures, as well as for the best sampled surface H atom interactions. To investigate the relevant energy surface for the MD simulations using experimental conditions, the system can be explored by ab-initio MD simulations. They are an expensive method, which probes a specific pathway through the PES very accurately, but due to its computational cost, is very limited in exploring large amounts of space on the PES. Picking the right areas to compare the prediction of the HDNNP with the reference method is difficult. Using the data generated by the HDNNP, MD simulation trajectories with different kinetic energy loss at different scattering angles can be picked to probe different scattering trajectories more easily. In the following section the HDNNP will be validated using the mentioned methods. In particular, comparing the visualization of the PES spanned by the grid structures and the prediction of the HDNNP with ab-initio MD simulations.

Visualization and Validation of the Potential using the H Atom Grid

The H atom grid used for the initial structure generation allows the identification of repulsive and attractive areas of the PES for the H atom. A good understanding of attractive and repulsive configurations can greatly improve the understanding of the system. It can reveal adsorption sides and diffusion channels, as well as point out the heights and structures of the surface visible for the H atom at different kinetic energies. In figure 4.13, eight 2D cuts through the PES in plane with the surface are shown. The points are interpolated using linear interpolation between points on the triangular grid. Figure 4.13 depicts three panels, with each panel representing a different type of the PES. Panel **(a)** displays the PES calculated using RPBE DFT. In panel **(b)**, the PES is predicted using the HDNNP. Finally, in panel **(c)** shows the difference between the DFT PES and the HDNNP prediction, highlighting discrepancies between the two models. The panels show the PES for the (2×2) super cell, the energy difference ΔE is calculated between the surface and H atom at infinite distance and the surface with the H atom at the distance z from the O layer. There are as many contour lines for energy differences below 0 and above 0 for the HDNNP and RPBE PES in panel **(a)** and **(b)**. This makes the structure of the PES in the attractive region of the PES visible, but also aggravates smaller differences between the HDNNP and the DFT data points. The symmetry of the surface is clearly represented by the PES. The attractive potential of the Al atom on top of the highest O-layer is visible in the middle of the cell, starting from the top. This minimum also depicts the threefold rotation axis, as well as being the global minimum for the H atom on top of the relaxed surface. Closer to the surface, the Al atom becomes repulsive for the H atom, drastically increasing the energy compared to the surrounding area covered by O atoms. The area around the Al atom, which is flanked by O atoms, is most attractive for the H atom at 3 Å off the surface, slowly becoming more repulsive toward the Al atom. At approximately 1.5 Å the O covered area becomes less repulsive than the Al atom. The H atom favors the area around the Al atom and will adsorb there under normal circumstances. The HDNNP potential in the middle panel of figure 4.13 reproduces all those features with high accuracy. The maximum deviation shown in the right panel is approximately -0.04 eV. This would represent an error of 0.3 meV atom^{-1}. Besides the maximum deviation at around 2 Å from the surface, the error is distributed over the surface, both with higher and lower predictions of the HDNNP. There is no bias toward higher or lower energies. Overall the prediction of the HDNNP accurately describes the PES for the H atom on top of the equilibrated surface, reproducing all features of the DFT PES.

The MD trajectories using experimental conditions that are calculated with the HDNNP have H atoms scattering from the surface with an initial kinetic energy of 1.92 eV. The HDNNP needs to be able to accurately describe the PES up to and beyond this threshold, since at the scattering point most of the kinetic energy will be converted to

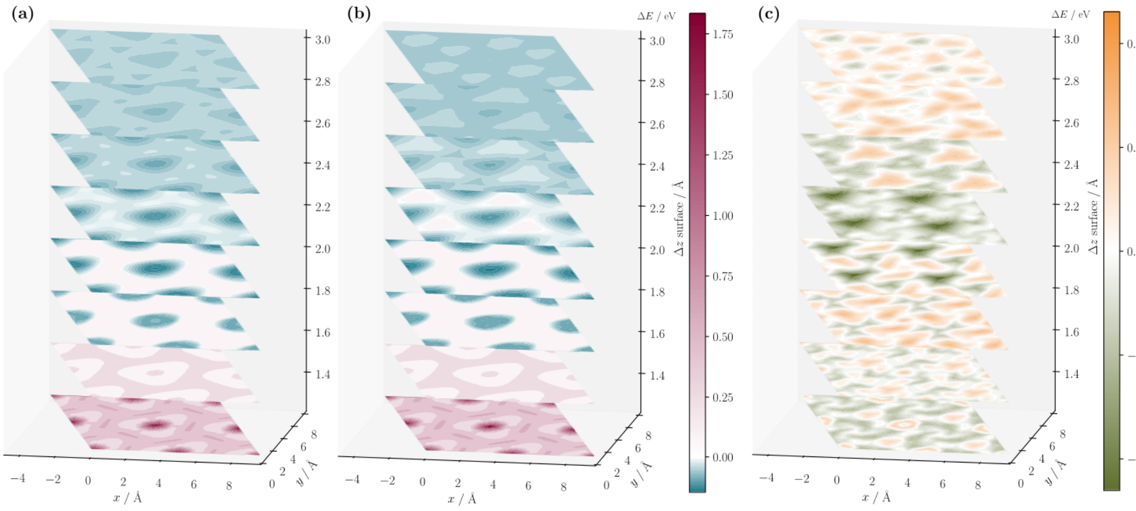


Fig. 4.13: 2D energy cuts through the PES of the H atom above the equilibrated surface. **(a)** depicts the RPBE PES, **(b)** shows the HDNNP PES and **(c)** shows the difference between both methods. The energy difference ΔE for the HDNNP and DFT energies is the difference between the optimized slab structure with the H atom at infinite distances compared to the H atom at the distance z from the O layer of the slab. The DFT data points are from the grid approach described in the construction of the data set in section 3.3.1, and the points are interpolated linearly on a trigonal grid.

potential energy. In order to better analyze the accuracy at high potential energies, 1D energy curves following the H atom coming closer to the surface can be plotted. In figure 4.14 the 1D energy curve for the H atom position above the topmost Al atom is shown up to an energy of 2.5 eV above the energy of the separated systems. Figure 4.15 shows the 1D energy curve for a H atom position above a O atom, as shown in the inset of the figure.

Both, the energy and force compontents in z direction of the H atom, are reproduced by the HDNNP up to at least 2.5 eV above the separated system. Small deviations at medium distances, better visible in figure 4.13, can be seen for both the energy and the force component. At large distances, the HDNNP accurately describes the separated system. The HDNNP can also accuratly describe the repulsive potential above the Al atom in figure 4.14, as well as the repulsive potential closer to the surface in figure 4.15.

The HDNNP exhibits high accuracy for grid structures overall, as demonstrated by the RMSE of the fit. Equilibrated slab structures are predicted with precision across all relevant energy ranges and H atom force components are replicated with similar precision. This comparison demonstrates a best-case scenario in which all relevant data points are present. However, in the upcoming MD trajectories at 300 K, not all possible slab geometries have a closely situated H atom grid. To assess the HDNNP's perfor-

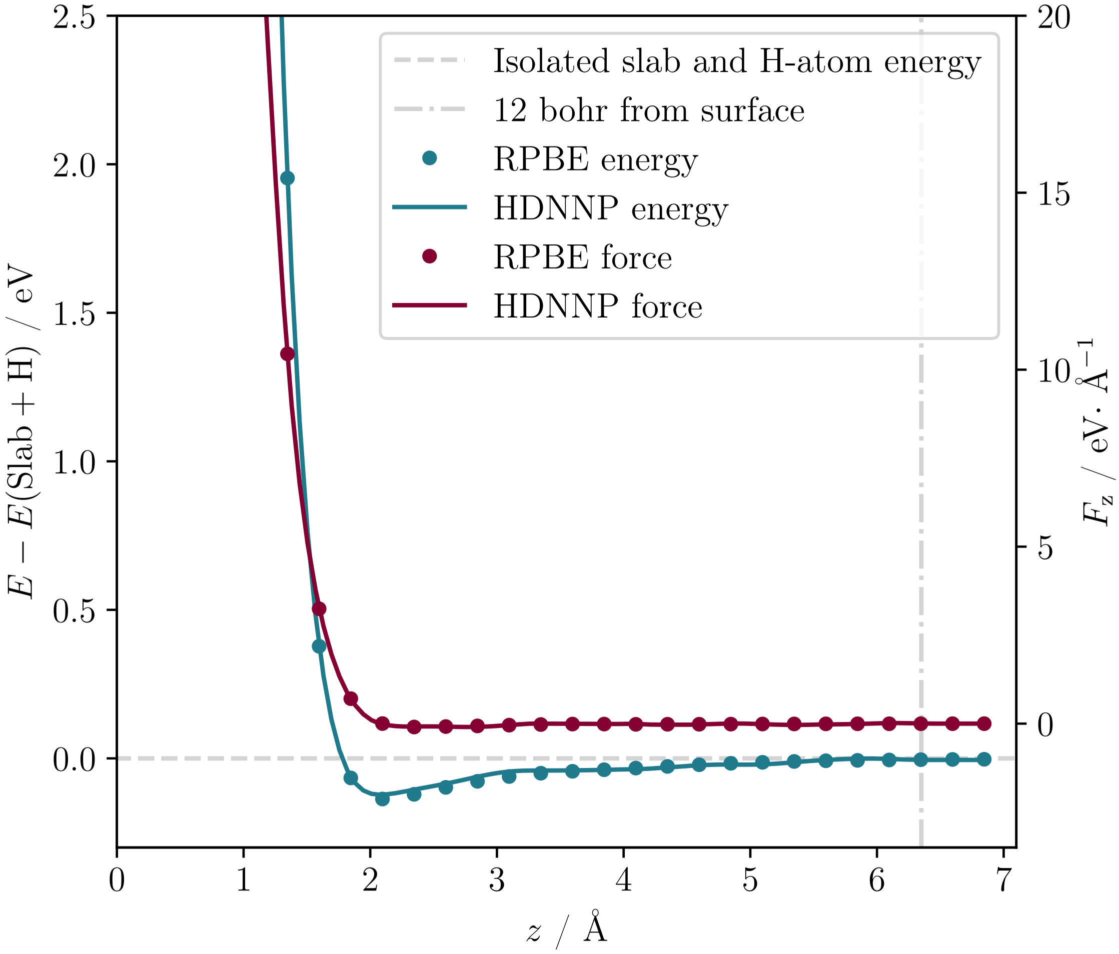


Fig. 4.14: 1D energy curve of H atom adsorbing on the equilibrated surface. The left axis shows the energy difference between the structure with the H atom at the distance z and the H atom at a infinite distance, in blue. The right axis shows the z direction force component of the H atom, in red. The grid DFT points are shown as circles, while the HDNNP points are calculated on a denser grid and shown as a line. The inset picture of the surface shows the position of the H atom in white above the surface.

mance on those structures, ab-initio trajectories are recalculated using the HDNNP in the next section.

Validation of the Potential using ab-initio Molecular Dynamic Simulations

Reference calculations are expensive and time consuming, the HDNNP can not be validated for every possible geometry. Usually HDNNP are applied for systems, which are difficult to describe using DFT calculations, because the system needs to be a certain size or the simulations need to be calculated for a long time. This will make the validation of the potential with those structures difficult. Here, the problem lies in

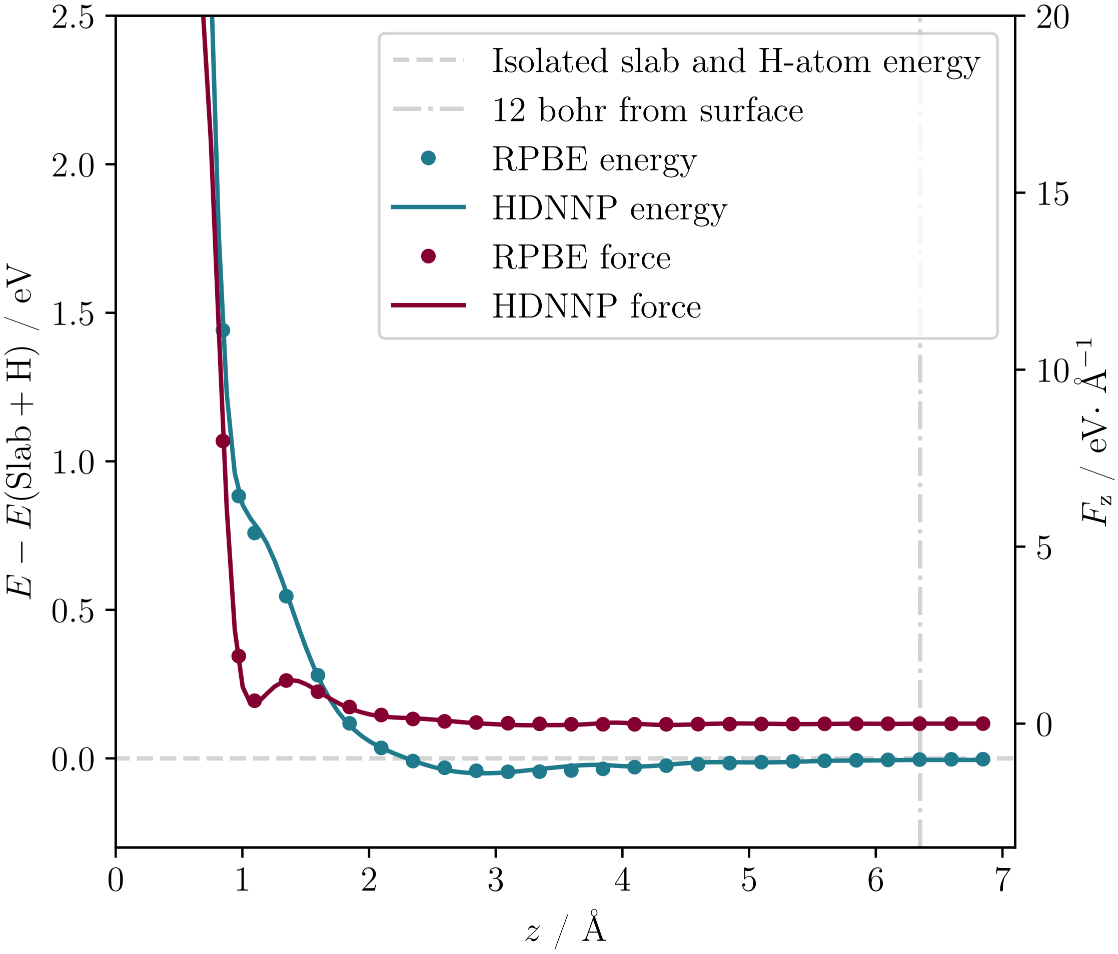


Fig. 4.15: 1D energy curve of H atom adsorbing on the equilibrated surface. The left axis shows the energy difference between the structure with the H atom at the distance z and the H atom at a infinite distance, in blue. The right axis shows the z direction force component of the H atom, in red. The grid DFT points are shown as circles, while the HDNNP points are calculated on a denser grid and shown as a line. The inset picture of the surface shows the position of the H atom in white above the surface.

the statistical nature of the experimental measurements. In the experiment, trillions of H atoms are scattered from the surface, all independently probing the surface and scattering, resulting in different scattering kinetic energies. Only those H atom that precisely hit the detector above the surface will be measured. Therefore, only a fraction of the simulations will fulfill the criteria needed to hit the detector and be plotted in the kinetic energy loss distribution. Trajectories hitting the detector at different scattering angles and energies can be directly compared with ab-initio trajectories to get the best overview of the accuracy of the HDNNP predictions as possible. Using the HDNNP to generate those trajectories at different points of the PES reduces the cost of generating ab-inito MD simulations. Running an ab-initio MD at the target temperature usually needs hundred of time steps, in order to equilibrate the structure to the right temperature. By starting the ab-inito MD from positions and velocities generated by the HDNNP, the cost can be reduced. The ab-initio trajectories, calculated to validate the performance of the HDNNP, were chosen from MD simulations using experimental conditions, calculated with a HDNNP potential. In total ten of those simulations were chosen, at different incident and scattering conditions shown in table 4.2. Those calculations were initially picked to represent different regions of the kinetic energy spectrum, especially to compare regions with good agreement between experiment and theory and regions with bad agreement. Here, one of the difficulties of describing the system becomes apparent. A set of incident conditions determining the starting position and velocities of all atoms can be picked based on the trajectory calculated with a HDNNP potential, resulting in certain scattering angles and kinetic energies. However, using those incident conditions, the ab-initio MD simulation does not have to result in similar scattering conditions. The chaotic nature of the MD simulation, where the small differences in the PES lead to larger differences in the trajectory, resulting in very different outcomes. This uncertainty means that it is unknown whether the selected incident conditions will result in similar scattering angles and kinetic energies as those of the original trajectory. This increases the difficulty in picking trajectories from different parts of the kinetic energy loss spectrum. It also means it is very difficult to compare the kinetic energy loss of a HDNNP trajectory with the kinetic energy loss of ab-initio MD simulations, where small differences in scattering angles could result in comparing two different trajectories. Therefore, in order to validate the HDNNP, the energies and forces of the structures of the ab-initio trajectories will be compared to prediction of the HDNNP of these structures.

Table 4.2 shows the incident and scattering conditions, as well as the RMSE for both energy and force of the H atom. The RMSE values are gathered by calculating every structure of the ab-inito trajectory with the HDNNP. The RMSE values for the potential energy and the force are smaller or very close to the RMSE values gathered from the training and test set. A similar RMSE indicates that the structures in the simulation are as well described as the structures in the data set. The data set seems to be well suited to describe the chemical environments probed by the ab-inito MD simulations. A closer look at two of the ab-initio MD simulations will reveal if the errors are spread over the whole trajectory or if they accumulate at the scattering incident.

The ab-initio trajectory MD4 (see Figure 4.16) shows a very simple scattering process, were the H atom follows a nearly right-angled trajectory scattering from a O atom. The z direction force component has one peak at the scattering incident at around $t = 40\,\text{fs}$. The potential energy of the system shows no large difference between before and after the scattering incident, since the H atom keeps most of its initial kinetic energy, only transferring around $0.14\,\text{eV}$. The energy errors are distributed along the

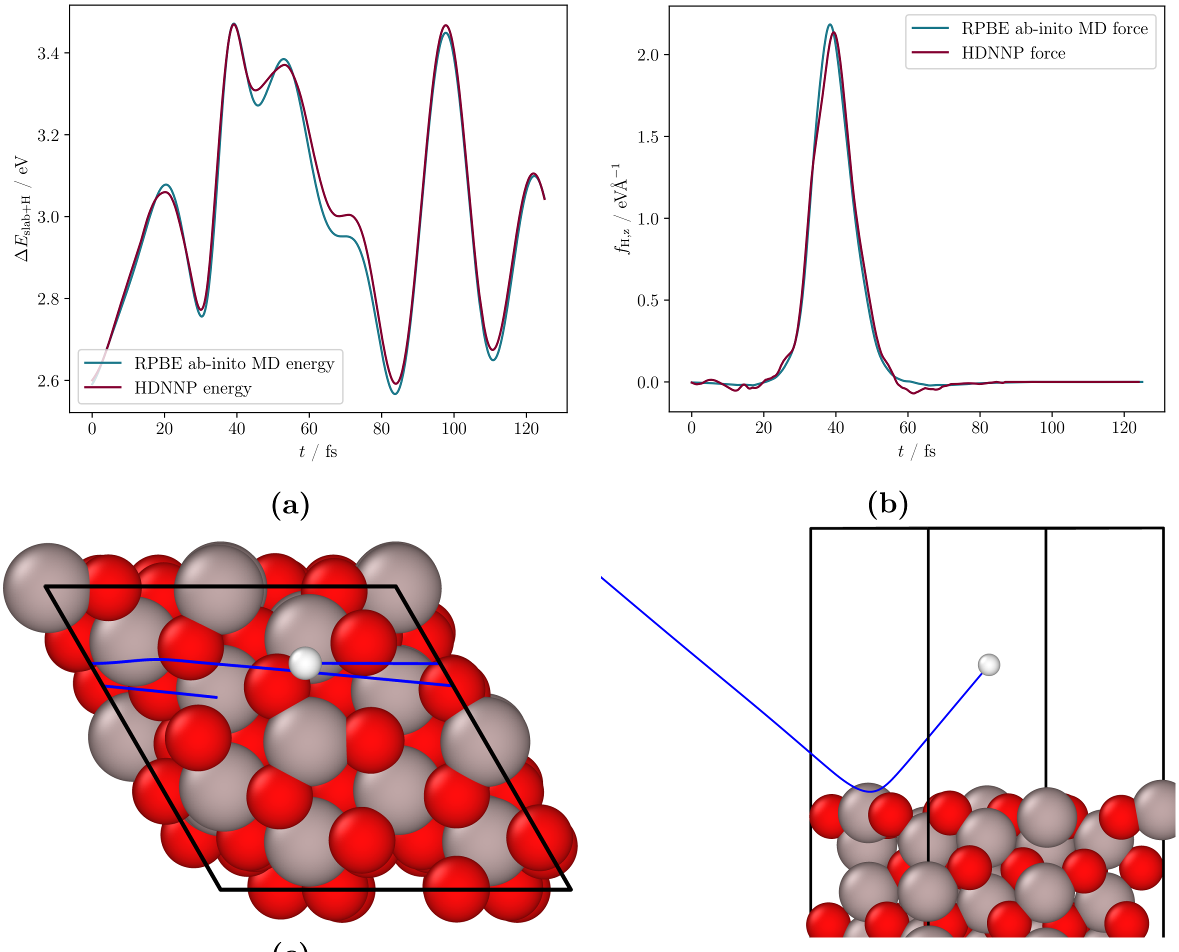
Tab. 4.2: Initial and scattering conditions of the ab-initio MD simulations used to validate the potential, as well as the RMSE for the energy and H atom force for the HDNNP prediction.

Name	Θ_i / °	ϕ_i / °	$E_\mathrm{kin,i}$ / eV	Θ_s / °	$E_\mathrm{kin,s}$ / eV	RMSE E / meV atom^{-1}	RMSE $F_\mathrm{H,z}$ / eV Å^{-1}
MD1	40	0	0.99	54	0.56	0.39	0.039
MD2	40	0	1.92	25	1.65	0.35	0.041
MD3	40	0	1.92	44	1.69	0.32	0.037
MD4	40	0	1.92	50	1.78	0.38	0.045
MD5	40	180	1.92	-10	1.00	0.34	0.125
MD6	40	180	1.92	32	1.00	0.33	0.047
MD7	40	180	1.92	61	0.83	0.35	0.102
MD8	40	180	1.92	13	0.57	0.47	0.108
MD9	55	0	0.99	49	0.83	0.36	0.022
MD10	55	180	1.92	48	1.70	0.37	0.047

whole trajectory, while the force error shows slightly more fluctuations at very small forces, compared to the peak force. Overall, there is a very good agreement between the ab-initio trajectory and the HDNNP potential.

The ab-initio trajectory MD5, shown in figure 4.16, shows a back scattering path, where the direction of the H atom is changed. Here, there is a clear difference between the potential energy before and after the scattering incident at $t = 40\,\mathrm{fs}$, because the H atom is losing 0.92 eV of kinetic energy. This higher kinetic energy loss is accommodated with a higher peak force of around 7.8 eV Å^{-1}, which is accurately predicted by the HDNNP. Overall, the HDNNP predicts the energy and force along the trajectory accurately.

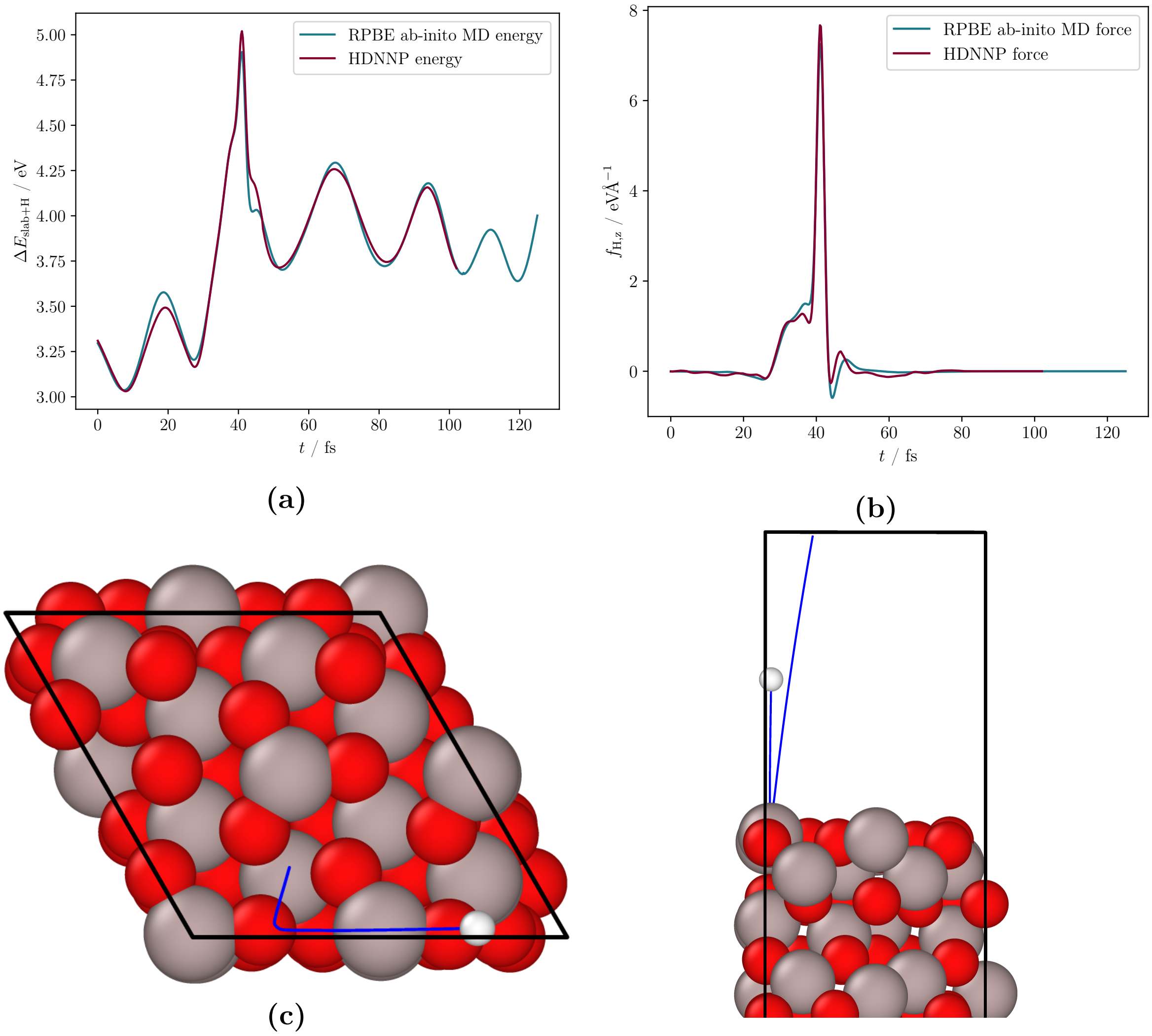
MD4 and MD5 are examples for trajectories with a high and with a small kinetic energy loss. The remaining ab-initio MD simulations show similar deviations for both energy and force components, but the amount of trajectoires is still too small for a statistical evaluation. The comparison between the ab-initio MD simulations and the prediction of the HDNNP shows that the high accuracy shown for the H atom grid and RMSE values is also expected for those unknown MD structures. It would also be possible to start HDNNP MD simulations with the same starting conditions as used for the ab-inito MD simulation, but due to the nature of MD simulations the smallest differences in the potential will lead to very different scattering conditions. This makes the comparison of the resulting kinetic energy loss impossible. Nonetheless the ability of the HDNNP to accurately predict the unknown structures of the ab-inito MD simulations shows that the HDNNP can accurately predict the full range of structures during a MD simulation using experimental conditions. From the high



(a) **(b)** **(c)** **(d)**

Fig. 4.16: Analysis of MD4. Panel **(a)** shows the potential energy of the structure $\Delta E_{\mathrm{slab+H}}$, in reference to the equilibrated slab with the H atom at infinite distance, vs. the time t of the trajectory, while panel **(b)** shows the z-direction force component of the H atom during the same time. Panel **(c)** and **(d)** show the path (blue) of the H atom (white) along the surface. In panel **(c)** the trajectory wraps at the edge of the cell. The surface is shown for the first frame of the simulation.

energy structures at the moment of scattering to the slab structure with higher kinetic energy after the scattering the prediction of the HDNNP is accurate. This lays the basis for comparing the experimental kinetic energy loss spectra with the results generated by the HDNNP in the next section.



(a) **(b)** **(c)** **(d)**

Fig. 4.17: Analysis of MD5. Panel **(a)** shows the potential energy of the structure $\Delta E_{\mathrm{slab+H}}$, in reference to the equilibrated slab with the H atom at infinite distance, vs. the time t of the trajectory, while panel **(b)** shows the z-direction force component of the H atom during the same time. Panel **(c)** and **(d)** show the path (blue) of the H atom (white) along the surface. The surface is shown for the first frame of the simulation.

4.2 Benchmarking the HDNNP with Experimental Data

4.2.1 Overview

Experimental benchmarks are needed to validate results of theoretical chemistry and provide a means of testing and refining theoretical models, allowing researchers to

gain a deeper understanding of chemical systems and make more accurate predictions. they also introduce challenges, since it is not always easy or possible to simulate an experimental setup accurately using theory. For good experimental benchmarks, the experiment has to be well defined, all conditions need to be known and the system cannot be too complicated so that calculations are still feasible to be carried out. The system of scattering H atoms from α-Al_2O_3 fulfills most of these characteristics. The ultra high vacuum conditions during the experiment create an environment were the surface should be clean from adsorbates and the H atoms should only interact with the surface. Additionally, the crystal is cleaned by argon sputtering and annealed at 700 °C to form a flat surface. The kinetic energy of the H atom and the beam are well defined and the resulting spectra show little noise. These are ideal conditions for theoretical benchmarking.
In this section, the creation of the theoretical kinetic energy distributions will be explained in detail. Afterwards the experimental kinetic energy distributions, as well as the angular distributions will be compared with the theoretical distributions. The system will be further analyzed in order to explore possible sources of deviations.

4.2.2 Creation of Theoretical Kinetic Energy Distributions

The experimental kinetic energy spectrum consist of a trillion separate H atoms scattered from the surface and detected by the detector. Every H atom has only one chance to hit the detector and for every scattering angle a separate set of scattered H atoms has to be measured. Here, for every incident condition, incident polar angle Θ_i and incident kinetic energy $E_\mathrm{kin.i}$, a separate data set consisting of hundred thousands of simulations has to be calculated. The setup for the MD simulations using experimental conditions is described in section 3.2. For the creation of the kinetic energy distribution and angular distribution spectra, the incident conditions of the H atom and the position and velocity at the end of the trajectory are used to evaluate if the simulation will be counted toward the respective distribution. If the H atom is on the way away from the surface, at a distance of at least 7.8 Å, it is eligible to be counted. For the H atom to hit the detector, the velocity vector $\vec{v}_\mathrm{H,s}$ has to point toward the detector. The detector position can be described by a unit vector $\hat{d}_\mathrm{detector}$. The relation between the x and y component of the $\hat{d}_\mathrm{detector}$ and the z component can be determined from the experimental detector polar angle Θ_det as

$$z = \cos\Theta_\mathrm{det} \ , \tag{4.6}$$

$$\sqrt{x^2+y^2} = \sin\Theta_\mathrm{det} \ , \tag{4.7}$$

and the relation between the x and y component can be determined by the experimental detector azimuth angle ϕ_{det} as

$$x = \sin\Theta_{\text{det}} \cdot \cos\phi_{\text{det}} \ , \tag{4.8}$$

$$y = \sin\Theta_{\text{det}} \cdot \sin\phi_{\text{det}} \ . \tag{4.9}$$

Combining equations 3.6 -3.9, $\hat{d}_{\text{detector}}$ can be calculated by the following equation:

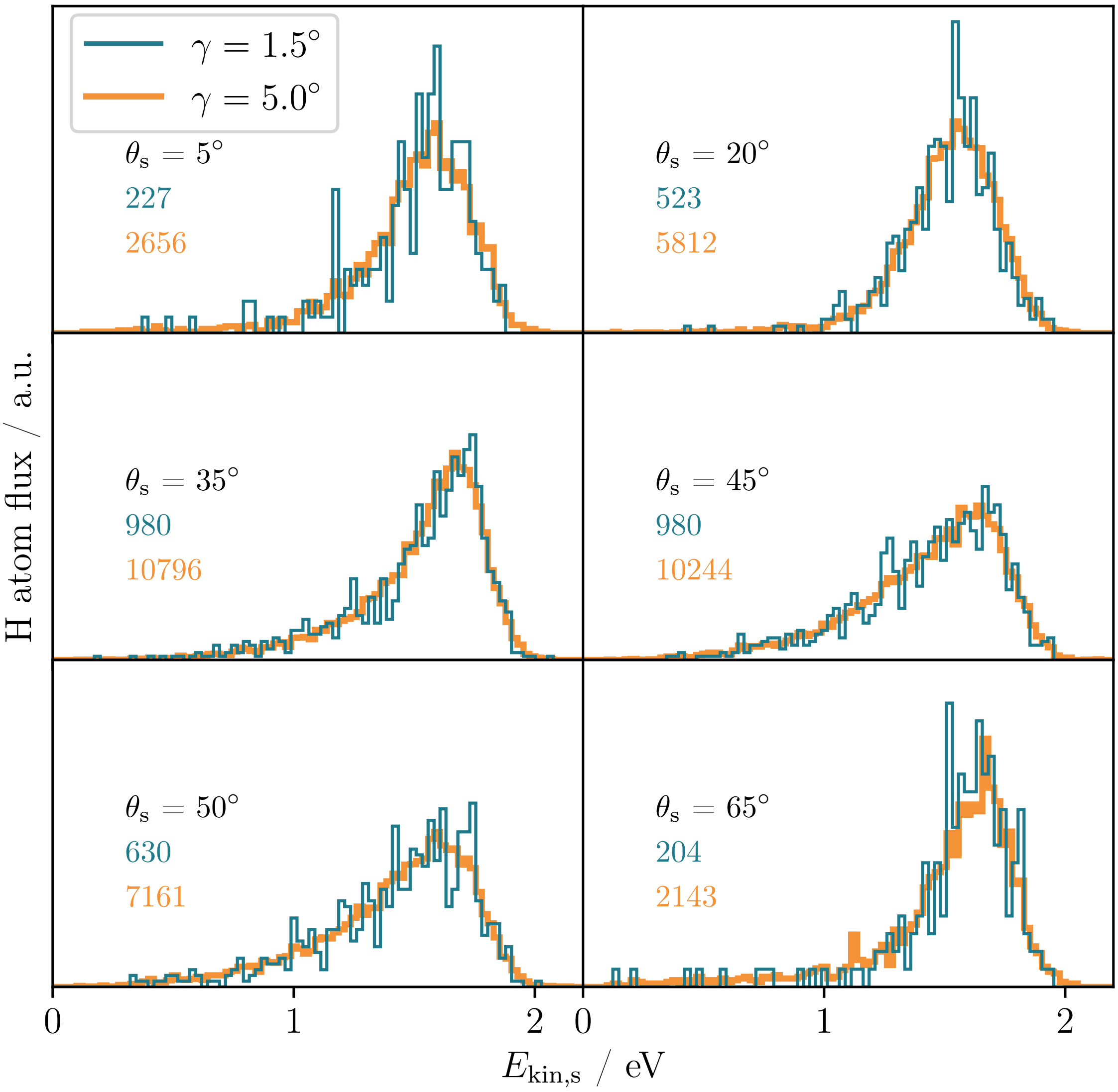
$$\hat{d}_{\text{detector}} = \begin{pmatrix} \sin\Theta_{\text{det}} \cdot \cos\phi_{\text{det}} \\ \sin\Theta_{\text{det}} \cdot \sin\phi_{\text{det}} \\ \cos\Theta_{\text{det}} \end{pmatrix}$$

The H atom velocity vector at the end of the trajectory $\vec{v}_{\text{H,s}}$, and its unit vector $\hat{v}_{\text{H,s}}$, can be used to calculate the angular difference γ (see equation 4.10) between the H atom velocity vector $\hat{v}_{\text{H,s}}$ and the detector position unit vector $\hat{d}_{\text{detector}}$.

$$\gamma = \cos^{-1}\left(\hat{v}_{\text{H,s}} \cdot \hat{d}_{\text{detector}}\right) \tag{4.10}$$

In the experiment the maximum γ between both vectors is determined by the size of the detector and its distance from the surface, resulting in an angular resolution of 3°, which results in $\gamma = 1.5°$. For the theoretical kinetic energy distributions, the threshold for γ can be adjusted to decrease the amount of trajectories needed to obtain a well defined spectra. As long as the distribution does not change, the value for γ can be increased to reduce the number of calculations necessary. Setting $\gamma = 5°$ increases the amount of counted trajectories by a factor of 10, while not changing the distribution, as shown in figure 4.18. While there are less trajectories contributing to the distribution with $\gamma = 1.5°$, resulting in a higher degree of fluctuation and variability in the distribution, the overall shape is the same for the distribution with $\gamma = 5°$.

The data points for a distribution are then sorted into 90 bins along the energy range (0 to 2.2 eV for $E_{\text{kin,i}} = 1.92\,\text{eV}$ and 0 to 1.45 eV for $E_{\text{kin,i}} = 0.99\,\text{eV}$). Both the theoretical and experimental distributions are scaled to the maximum of the distribution so that emphasis is put on the shape of the distribution as well as the onset and width of the distribution. The distributions will be shown for the experimentally measured scattering polar angles.



(a)

Fig. 4.18: Kinetic energy distribution plots for $E_{\text{kin,i}} = 1.92\,\text{eV}$, $\Theta_\text{i} = 40°$ and $\phi_\text{i} = 0°$. The orange line in the background shows the distribution for $\gamma = 5°$ and the blue line for $\gamma = 5°$. The number of trajectories hitting the detector at the corresponding Θ_s is written in the plot, color coded for the respective γ.

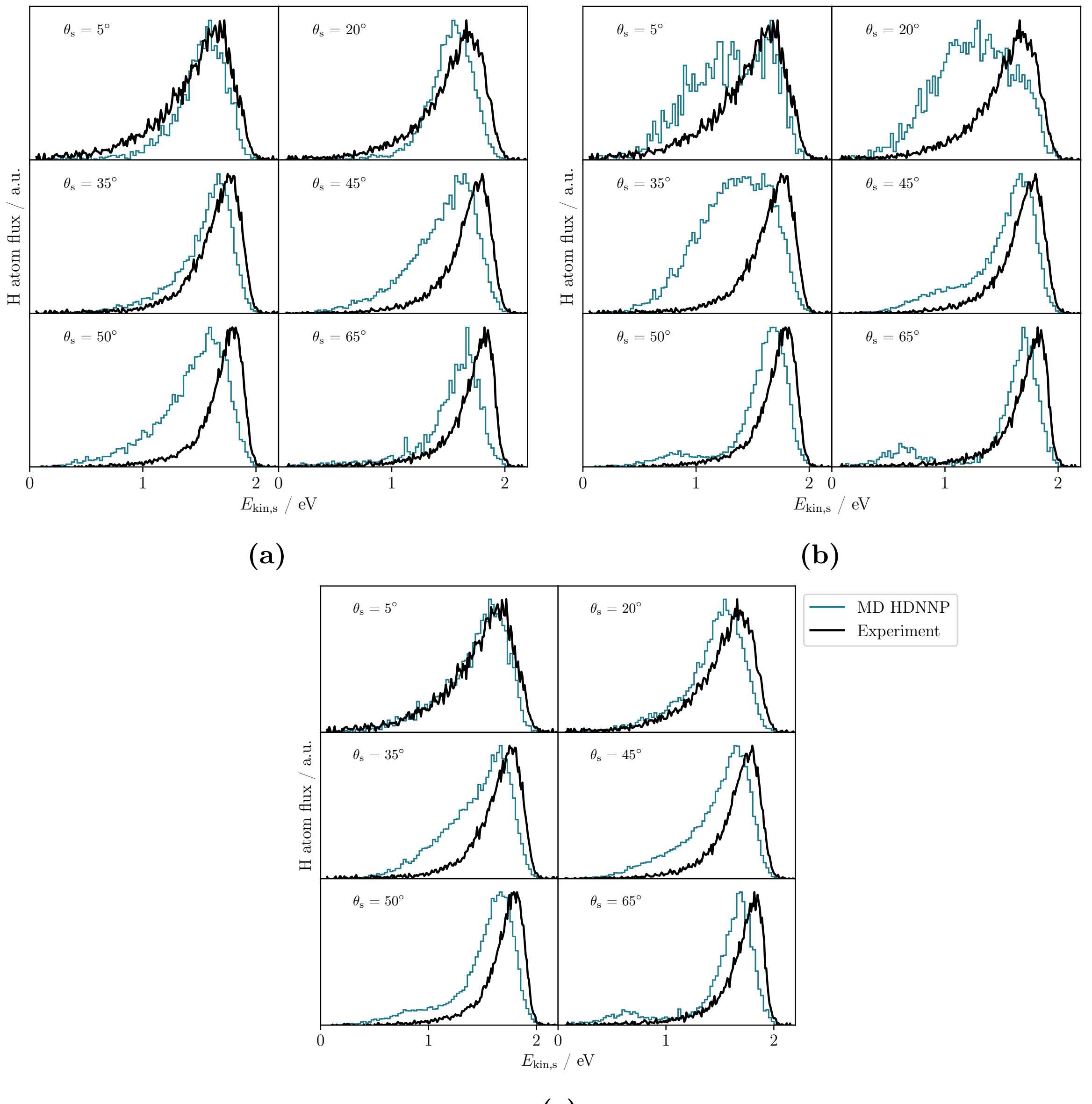
4.2.3 Benchmarking Experimental and Theoretical Kinetic Energy Distributions

The experimental data set consist of four separate sets of incident conditions, $E_{\text{kin,i}} = 0.99\,\text{eV}$ with $\Theta_\text{i} = 40°$, $E_{\text{kin,i}} = 0.99\,\text{eV}$ with $\Theta_\text{i} = 55°$, $E_{\text{kin,i}} = 1.92\,\text{eV}$ with $\Theta_\text{i} = 40°$ and $E_{\text{kin,i}} = 1.92\,\text{eV}$ with $\Theta_\text{i} = 55°$. For each set at $\Theta_\text{i} = 40°$, there are six Θ_s measured and for each set at $\Theta_\text{i} = 55°$ there are three Θ_s measured. For every incident condition set, there is also an angular distribution measured. The detector was put opposite of the H atom beam for all measurements, resulting in $\phi_\mathrm{i} = 0°$. The [0001] α-Al_2O_3 surface introduces another variable into the simulation, it can be terminated in two ways, which are chemically and energetically equivalent, but will produce significantly different scattering spectra. The top most oxygen layer can be replicated by the second oxygen layer via a mirror operation.[98] The crystal used during the experiment was not a perfect single crystal and thous should have some steps on the surface. Scattering on the two sides of the step will result in scattering at a incident azimuthal angle ϕ_i of either 0° or 180°, depending on the oxygen layering. There is no way to ensure that the experimental data was scattered from only one oxygen layering. Comparing the experimentally measured distributions with the distributions using the HDNNP, the oxygen layering can be predicted. It could be scattering from either layering or a combination of both.

In figure 4.19, the kinetic energy loss spectrum for E_kin =1.92 eV and $\Theta_\mathrm{i} = 40°$ at ϕ_i =0° **(a)** and ϕ_i =180° **(b)**, as well as the combination of both spectra **(c)** is compared to the experimental spectra.

The experimental kinetic energy loss distributions shown in all three plots in figure 4.19 are the same. For every set of incident conditions, around 300000 MD simulations using the HDNNP have been calculated. The kinetic energy distributions narrow moving from small scattering polar angle toward larger angles, while the maximum of the distribution shifts toward higher kinetic energies, in accordance with the binary collision model. The theoretical distributions for $\phi_\mathrm{i} = 0°$ shows a generally lower kinetic energy at the onset and maximum of the distribution, while the distributions at $\Theta_\mathrm{s} = 5°$, 20°, 35° and 65° are narrower or as narrow as the experimental distributions and the distributions for $\Theta_\mathrm{s} = 45°$ and 50° are wider than the experimental ones while a keeping similar shape. The distributions for ϕ_i=180° shows a lower kinetic energy at the onset and maximum for Θ_s up to 35°, while the kinetic energy at larger Θ_s is closer to the experimental one. The distributions for $\Theta_\mathrm{s} = 45°$, 50° and 65° are very similar in width, shape and kinetic energy at the maximum and the the distributions at $\Theta_\mathrm{s} = 5°$, 20° and 35° are wider than the experimental distributions.

Comparing both sets of distributions to the experiment shows them matching better at different angles. At small scattering angles the distribution for ϕ_i=0° deviates less compared to the experiment, while at larger scattering angles the distribution of ϕ_i=180° fits better. The distribution using a combination of both scattering azimuth angles more uniformly agrees with the experiment. For $\Theta_\mathrm{s} = 5°$, the agreement between experiment and theory is almost perfect, while $\Theta_\mathrm{s} = 20°$ shows a slight shift toward lower kinetic energies. The distributions with $\Theta_\mathrm{s} = 35°$, 45° and 50° are all slightly wider and shifted toward lower kinetic energies, while $\Theta_\mathrm{s} = 35°$ has a slower decline of the distribution between 1.5 eV and 0.8 eV. The distribution with $\Theta_\mathrm{s} = 65°$ again matches well with the experiment with a slight shift toward lower kinetic energies.



(a) **(b)**

(c)

Fig. 4.19: Kinetic energy loss distributions for H atom scattering at $E_{\text{kin}} = 1.92\,\text{eV}$ and $\Theta_{\text{i}} = 40°$. The incident azimuthal angle ϕ_{i} is 0° for **(a)**, 180° for **(b)** and a combination of both for **(c)**, which is weighted by the amount of trajectories calculated for the incident condition. For every set of incident conditions, around 300000 MD simulations using the HDNNP have been calculated. The experimental and theoretical distributions are scaled to the maximum of the distribution.

The distributions for $E_{\text{kin}} = 0.99\,\text{eV}$ and $\Theta_{\text{i}} = 40°$ in figure 4.20 are all more similar in shape. While the shape of the experimental distribution narrows with increasing scattering angle, the width of the theoretical distribution does not change. Comparing the theoretical distributions for $\phi_{\text{i}} = 0°$ and $\phi_{\text{i}} = 180°$ there are no clear differences visible. All three theoretical sets of distributions, $\phi_\mathrm{i} = 0°$, $\phi_\mathrm{i} = 180°$ and the combination of both, are very similar in position and shape of the distributions. While the kinetic energy at the maximum of the distribution for E_kin =1.92 eV and $\Theta_\mathrm{i} = 40°$ increased with increasing Θ_s, the kinetic energy at the maximum of the distribution for E_kin =0.99 eV and $\Theta_\mathrm{i} = 40°$ does barely increases and decreases for scattering at $\Theta_\mathrm{s} = 65°$. The H atoms scattering at $\Theta_\mathrm{s} = 65°$ are staying in the attractive part of the surface the longest and might lose more kinetic energy in the process.

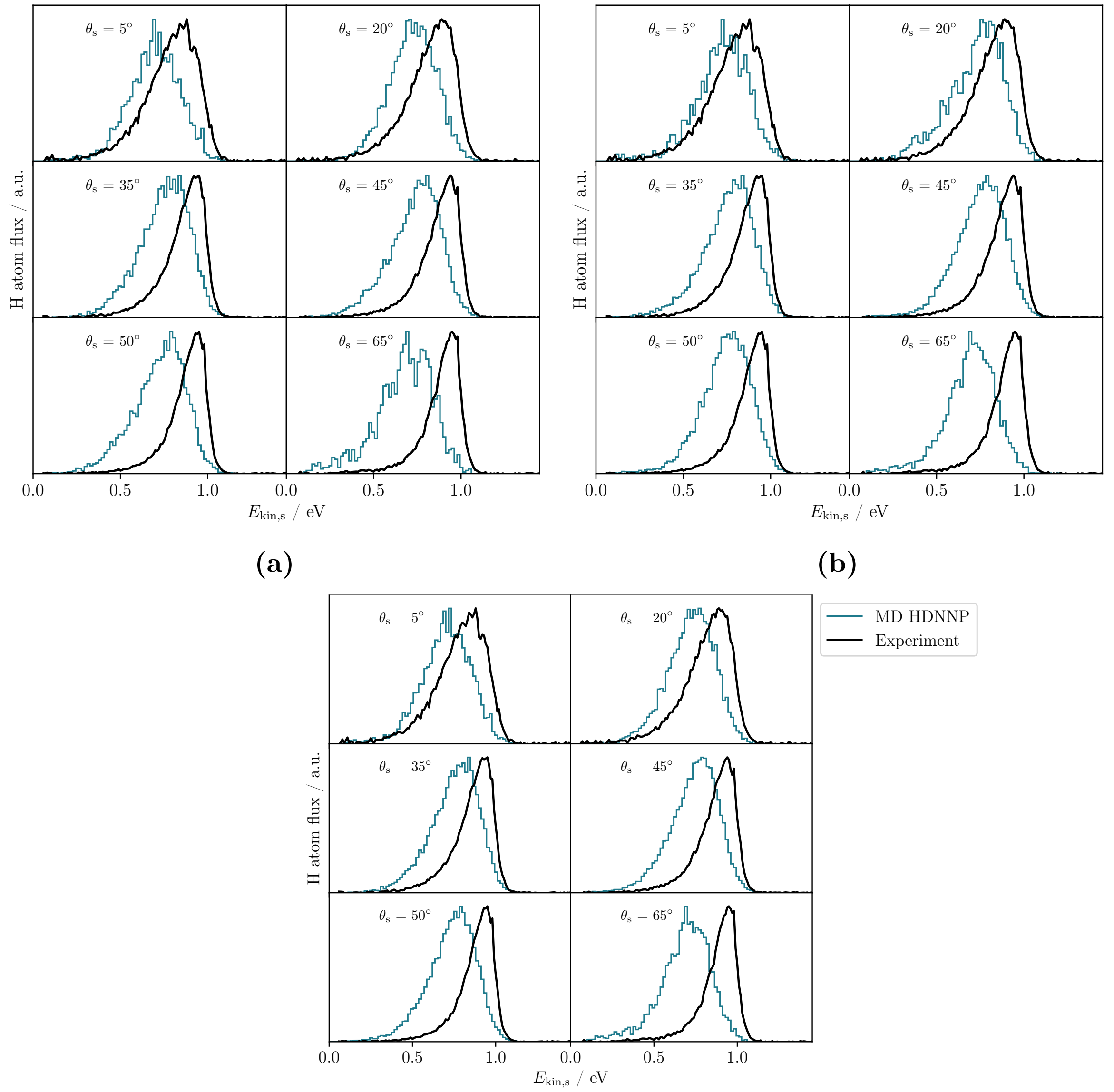
For the distributions with $\Theta_\mathrm{i} = 55°$, there are only experimental measurements for $\Theta_\mathrm{s} = 35°$, 55° and 65°. This is due to the fact, that at small scattering angles not enough H atoms are scattered. The theoretical distribution for $\Theta_\mathrm{s} = 5°$, even with increased detector size, is statistically not relevant, while for $\Theta_\mathrm{s} = 20°$ and 35° enough H atoms hit the detector, making the analysis of the distributions feasible.

The scattering at $E_\mathrm{kin} = 0.99$ eV and $\Theta_\mathrm{i} = 55°$, in figure 4.21, shows many similarities to the scattering at $E_\mathrm{kin} = 0.99$ eV and $\Theta_\mathrm{i} = 40°$. The distributions, except for $\Theta_\mathrm{s} = 5°$, all are of a similar shape. The maximum of the distribution changes again from lower toward higher kinetic energies when moving from small scattering angles toward the larger ones, except for the scattering at $\Theta_\mathrm{s} = 65°$, for which the maximum is at slightly lower kinetic energies.

The scattering at $E_\mathrm{kin} = 1.92$ eV and $\Theta_\mathrm{i} = 55°$, in figure 4.22, shows the differences between the scattering at $\phi_\mathrm{i} = 0°$ and $\phi_\mathrm{i} = 180°$ for scattering at $\Theta_\mathrm{s} = 20°$, 35° and 45°, where the distribution for $\phi_\mathrm{i} = 180°$ is broader and shows a shoulder toward lower kinetic energies, very similar to the scattering at $E_\mathrm{kin} = 1.92$ eV and $\Theta_\mathrm{i} = 40°$. The distribution for $\Theta_\mathrm{s} = 55°$ and 65° does not change between scattering at $\phi_\mathrm{i} = 0°$ and $\phi_\mathrm{i} = 180°$. Again the combination of both scattering azimuth angles reproduces the experimental distributions the best, especially for the scattering at $\Theta_\mathrm{s} = 35°$, where the shape of the theoretical distribution is very similar to the experimental one.

Considering the available data, there is strong evidence toward scattering from both oxygen layers, giving a more coherent comparison between experiment and theory. The scattering with $E_\mathrm{kin} = 1.92$ eV is described best by the combination of both scattering azimuth angles. For both incident polar angles, the distributions at low scattering angles at $\phi_\mathrm{i} = 0°$ are narrower than the experimental one and show a broader distribution for scattering at $\phi_\mathrm{i} = 180°$. The scattering with $E_\mathrm{kin} = 0.99$ eV shows little to no difference between the distributions for $\phi_\mathrm{i} = 0°$ and $\phi_\mathrm{i} = 180°$, neither for low nor for high scattering polar angles.

The influence of the incident azimuth angle seems to be the strongest at high kinetic energies and low scattering polar angles. A higher kinetic energy enables the H atom to get closer toward surface, giving a more nuanced representation of the surface atom positions. The same is true for scattering at low polar angles, where the H atom has to be deflected upwards instead of grazing the surface as is the case for low scattering polar angles. From a solely geometric viewpoint the stronger dependence on the inci-



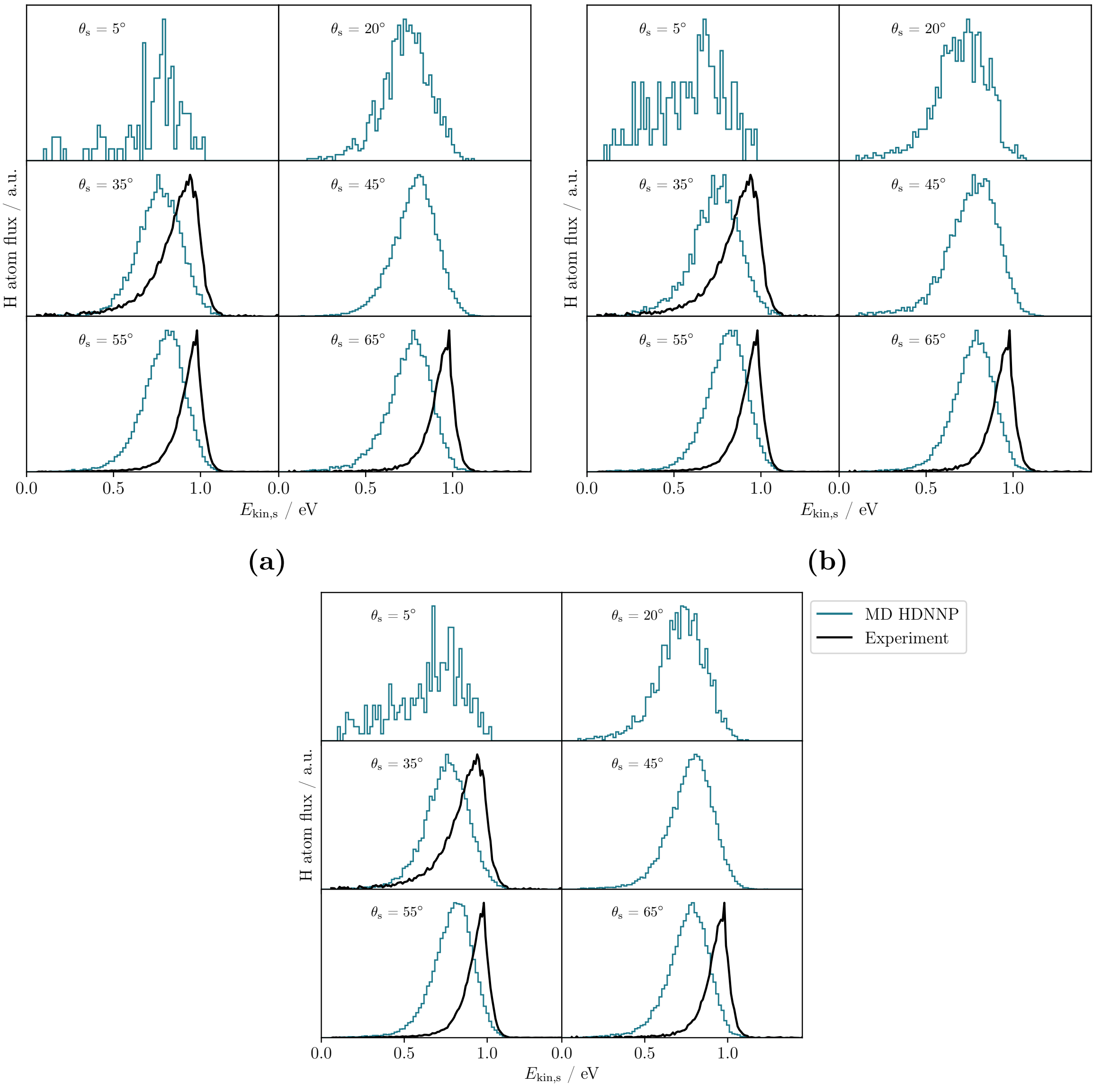
(a) **(b)**

(c)

Fig. 4.20: Kinetic energy loss distributions for H atom scattering at $E_{\text{kin}} = 0.99\,\text{eV}$ and $\Theta_{\text{i}} = 40°$. The incident azimuthal angle ϕ_{i} is 0° for **(a)**, 180° for **(b)** and a combination of both for **(c)**, which is weighted by the amount of trajectories calculated for the incident condition. For every set of incident conditions around 300000 MD simulations using the HDNNP have been calculated. The experimental and theoretical distributions are scaled to the maximum of the distribution.

dent azimuth angle for the $E_{\text{kin}} = 1.92\,\text{eV}$ and $\Theta_{\text{i}} = 40°$ case seems to be physically reasonable.

Considering the results discussed in this section, all following analysis will be done on the combination of the scattering at $\phi_{\text{i}} = 0°$ and $\phi_{\text{i}} = 180°$. The HDNNP accurately

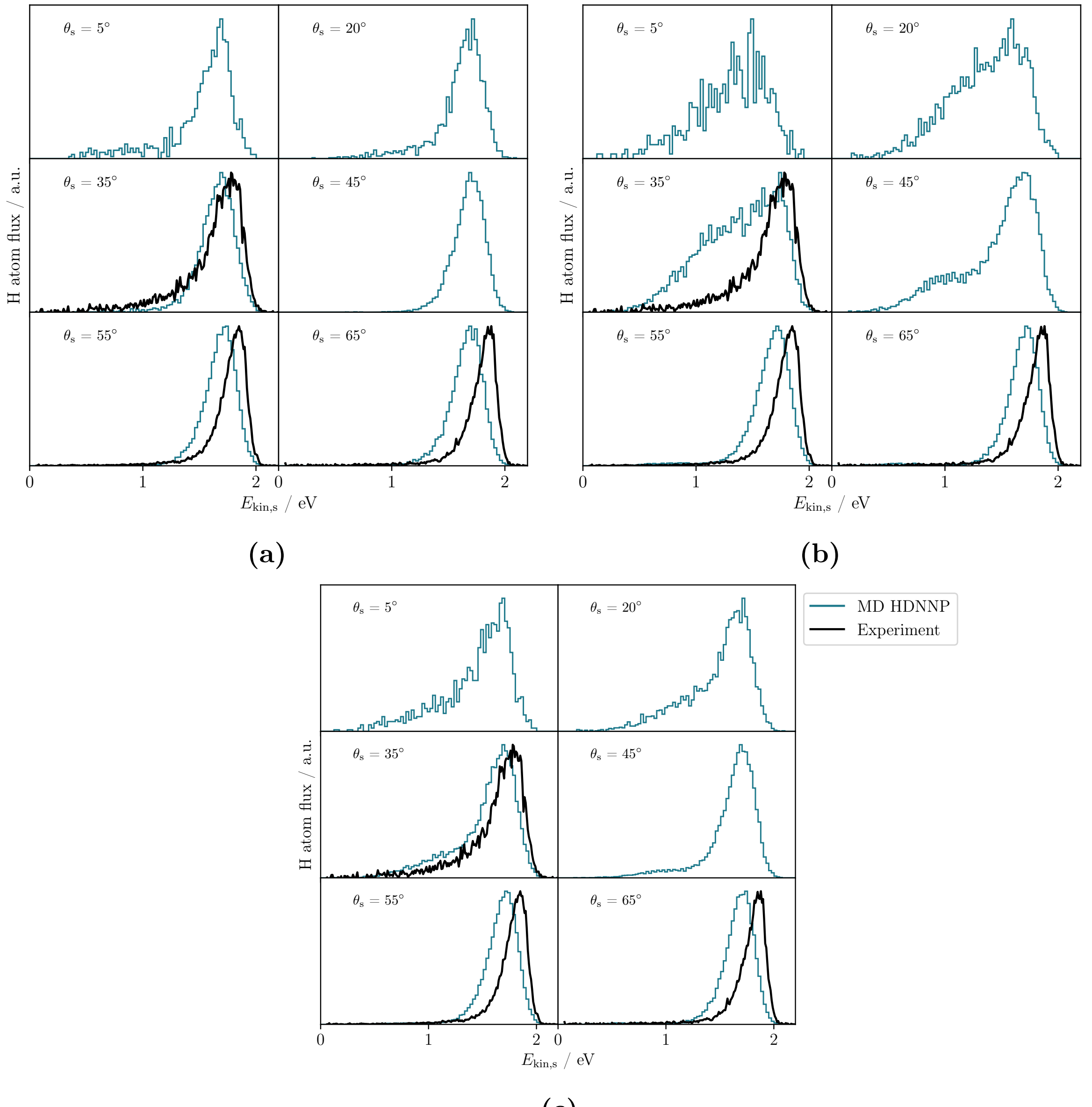


(a) **(b)**

(c)

Fig. 4.21: Kinetic energy loss distributions for H atom scattering at $E_{\text{kin}} = 0.99\,\text{eV}$ and $\Theta_i = 55°$. The incident azimuthal angle ϕ_i is 0° for **(a)**, 180° for **(b)** and a combination of both for **(c)**, which is weighted by the amount of trajectories calculated for the incident condition. For every set of incident conditions around 300000 MD simulations using the HDNNP have been calculated. The experimental and theoretical distributions are scaled to the maximum of the distribution.

predicts most of the features of the experimental distributions, especially the distributions for $E_{\text{kin}} = 1.92\,\text{eV}$ and $\Theta_i = 40°$ are well described in shape. All distributions show more kinetic energy loss than in the experiment, except for the scattering at $E_{\text{kin}} = 1.92\,\text{eV}$, $\Theta_i = 40°$ and $\Theta_s = 5°$. The scattering with $E_{\text{kin}} = 0.99\,\text{eV}$ has a larger



(a) **(b)** **(c)**

Fig. 4.22: Kinetic energy loss distributions for H atom scattering at $E_{\text{kin}} = 1.92\,\text{eV}$ and $\Theta_{\text{i}} = 55°$. The incident azimuthal angle ϕ_{i} is 0° for **(a)**, 180° for **(b)** and a combination of both for **(c)**, which is weighted by the amount of trajectories calculated for the incident condition. For every set of incident conditions around 300000 MD simulations using the HDNNP have been calculated. The experimental and theoretical distributions are scaled to the maximum of the distribution.

relative energy loss, especially for high scattering angles.

The general differences between the experimental and theoretical distributions and their likely causes will be discussed a following section, after gathering all the needed information. The main differences are the higher mean kinetic energy loss, which is also indicated by the maximum of the distribution, the mostly broader shape of the distributions and punctual differences, e.g. the extra peak for scattering at $E_{\mathrm{kin}} = 1.92\,\mathrm{eV}$, $\Theta_{\mathrm{i}} = 40°$ and $\Theta_{\mathrm{s}} = 65°$.

The discrete peak at low kinetic energies for scattering at $E_{\mathrm{kin,i}} = 1.92\,\mathrm{eV}$, $\Theta_{\mathrm{i}} = 40°$ and $\Theta_{\mathrm{s}} = 65°$ (see figure 4.19) needs to be investigated. It is a well defined peak resulting from the scattering with $\phi_{\mathrm{i}} = 180°$ which is not visible in the experimental spectrum. Since it is a well defined peak, it is expected to result from certain region of the surface, instead of being a fitting error, where the peak consists of H atoms scattered from random places of the surface. In order to investigate the scattering origin on the surface, a very simple approach will be applied. For each trajectory, the point at which the H atom is closest to the surface (referred to as the point of scattering in the following analysis) can be plotted. Since the surface is in motion during the simulation, a exact x and y position of the point of scattering is not very meaningful, since the surface will be different for each trajectory. With this methodology the origin of the scattering event can be approximated reasonably well allowing to pinpoint possible reasons for the deviations between the experiment and theory. The point of scattering will be defined as the lowest position of the H atom in the trajectory. A more complex definition of this point is not needed, since the goal is to generate an overview of the area of scattering. In figure 4.23 the point of scattering for all trajectories hitting the detector at $E_{\mathrm{kin,i}} = 1.92\,\mathrm{eV}$, $\Theta_{\mathrm{i}} = 40°$, $\Theta_{\mathrm{s}} = 65°$ and $E_{\mathrm{kin,s}} <= 0.9\,\mathrm{eV}$ is plotted over the unit cell of the surface.

All the scattering for that peak comes almost exclusively from one place on the surface. Examining the trajectories, it can be observed that they come into contact with the O atom before bending around the Al atom, transitioning from a more acute angle to a 65° angle, ultimately aiding in the formation of the peak. This means that the outcomes of the trajectories are caused by a localized feature of the PES, given that during multiple different MD simulations this scattering happened at the same place above the surface. Therefore, this scattering behavior is a feature of the fitted PES. In order to determine if it is also a feature of the RPBE PES, points along the trajectories can be calculated using DFT to determine the accuracy of the HDNNP for those trajectories.

Comparing the energy and force predictions of the HDNNP in figure 4.24 with the single point calculations using DFT shows no large differences. The trajectory in panel **(b)** deviates the most at the end of the DFT points, at around $0.1\,\mathrm{eV}$, while there are only small differences for the force prediction visible at the peak of the force in panel **(c)**. Overall, the accuracy of the simulations is very similar to the accuracy of the ab-initio trajectories and shows no substantial deviations that would hint toward a hole in the PES. While this does not exclude the possibility that the peak is a artifact of the HDNNP fit, it shows that the predicted energies and forces during these simulations are as accurate as expected from the prior analysis. Additionally, this feature was present for multiple generations of the HDNNP, further decreasing the probability of it

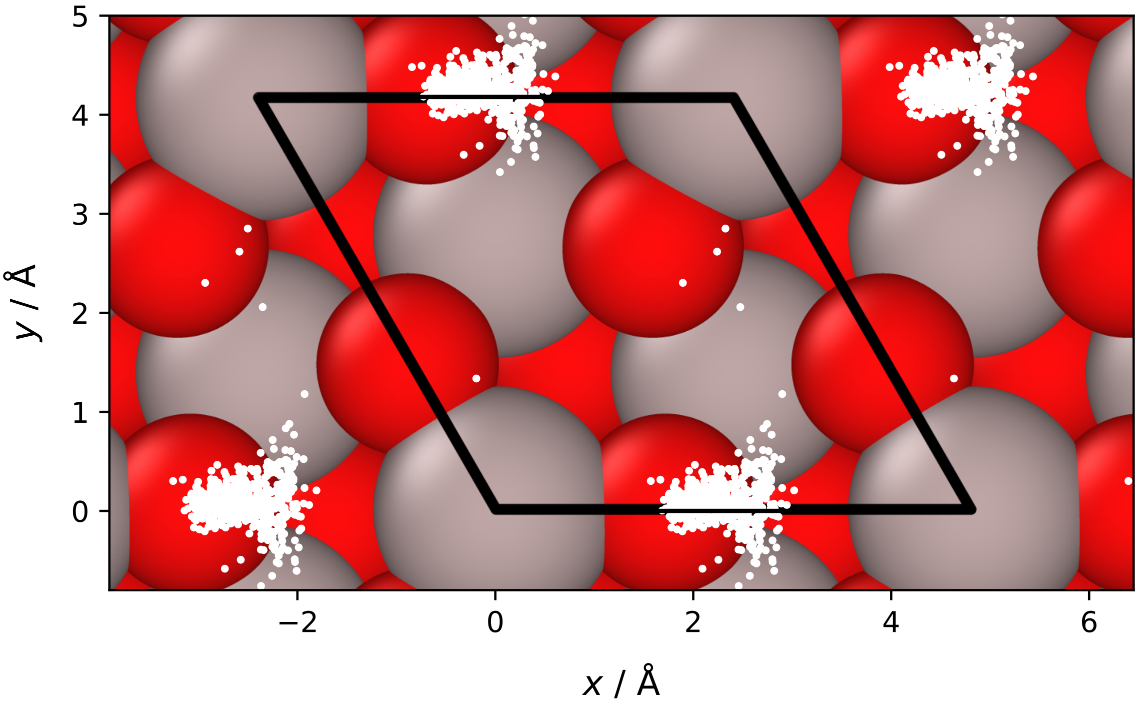


Fig. 4.23: x and y coordinate of each trajectory where the H atom gets closest to the surface on top of the surface picture for trajectories with $E_{\mathrm{kin,i}} = 1.92\,\mathrm{eV}$, $\Theta_{\mathrm{i}} = 40°$, $\Theta_{\mathrm{s}} = 65°$ and $E_{\mathrm{kin,s}} <= 0.9\,\mathrm{eV}$. The points come from trajectories with moving surfaces, they indicate the area from where the scattering is possible.

happening due to a hole in the PES. If it was due to a hole in the PES, the prediction of the HDNNP in that region should be random and not consistently lead to the same peak. With the gathered information it is very likely for this peak to be a feature of the RPBE PES and not due to inaccurate fitting of the HDNNP.
For the further analysis of the influence of the functional, the MD simulations and experimental conditions for all available data should be taken into account. The discussion of these points will be continued after the analysis of the angular distributions.

4.2.4 Benchmarking Angular Distributions

Additionally to the kinetic energy distributions, for each set of incident conditions the angular distribution of the scattered H atoms were measured experimentally. For the theoretical description of the angular distributions, the detector was placed in the same manner as for the kinetic energy distributions. However, instead of measuring just at the experimental scattering polar angle, the detector was placed at multiple positions between −10° and 90° in 3° steps. Then for each polar scattering angle the amount of trajectories hitting the detector at that angle, normalized to the maximum, was plotted vs the scattering angle. In figure 4.25, the angular distributions are shown for

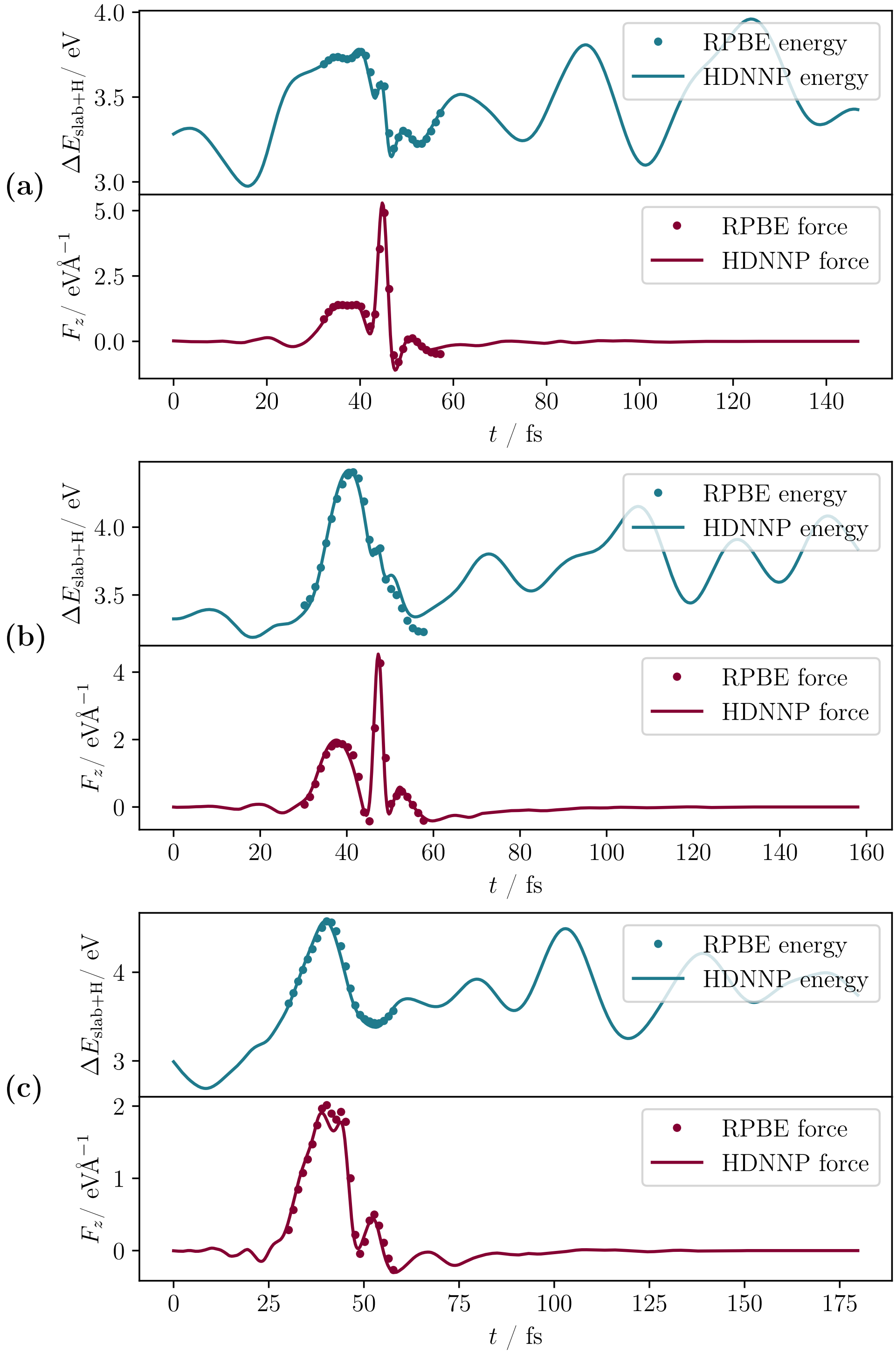


Fig. 4.24: Potential energy and z direction force component of three trajectories contributing to the low kinetic energy peak at $E_{\mathrm{kin}} = 1.92\,\mathrm{eV}$, $\Theta_{\mathrm{i}} = 40°$ and $\Theta_{\mathrm{s}} = 65°$. The upper part shows the potential energy relative to optimized slab and H atom at infinite distance, while the bottom part shows the z direction force component of the H atom vs the time of the simulation. Around the time of the collision 22 DFT points have been calculated to compare with the HDNNP energy and force.

all incident conditions.

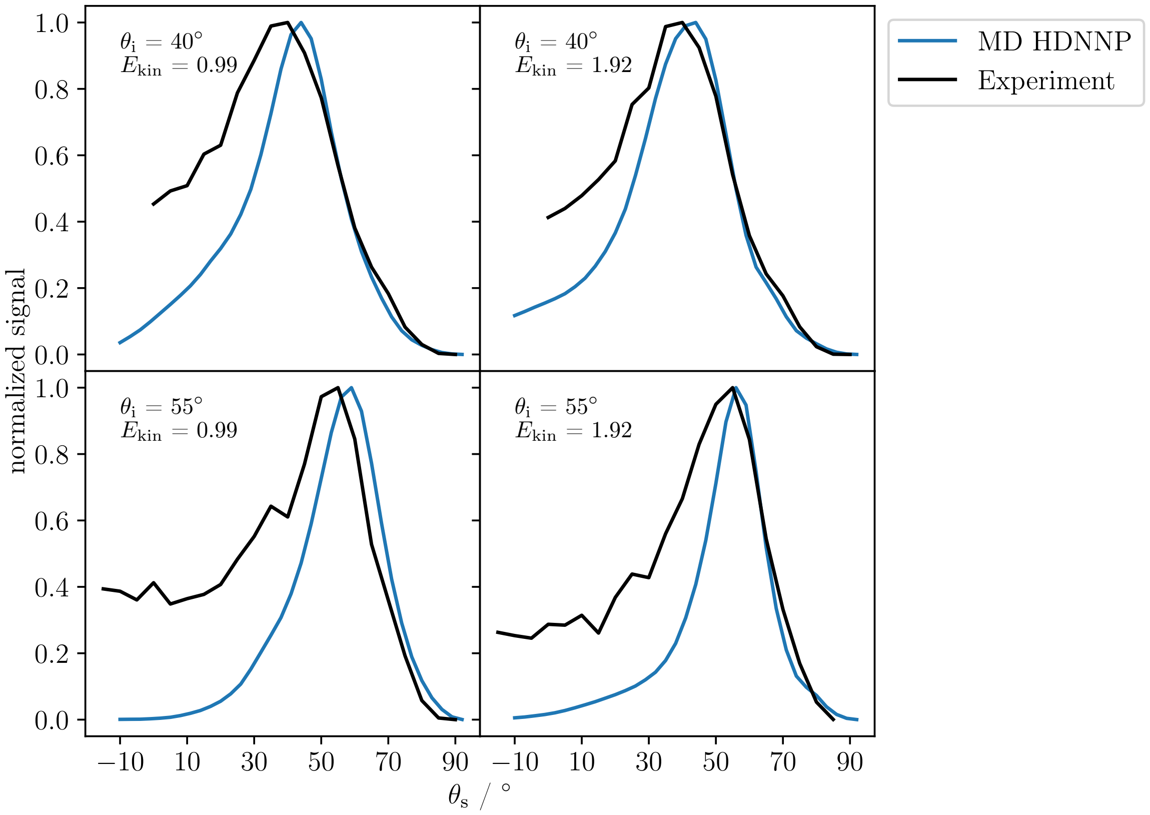


Fig. 4.25: Normalized angular distribution for scattering at (a) $E_{kin,i} = 0.99\,eV$ and $\Theta_i = 40°$, (b) $E_{kin,i} = 1.92\,eV$ and $\Theta_i = 40°$, (c) $E_{kin,i} = 0.99\,eV$ and $\Theta_i = 55°$ and (d) $E_{kin,i} = 1.92\,eV$ and $\Theta_i = 40°$. The blue lines are the theoretical distributions calculated using the HDNNP, while the black lines are the experimental distributions. Both theoretical and experimental distributions are normalized to 1 at the maximum.

Panel **(a)** shows the angular distribution for $E_{kin,i} = 0.99\,eV$ and $\Theta_i = 40°$. While the experimental distribution has its maximum at the specular angle of 40°, the maximum of the theoretical distribution is at 45°. Toward larger scattering angles the theoretical distributions agree well with the experiment, while toward scattering angles smaller than the specular angle the difference between the theoretical and experimental distribution grows. At $\Theta_s = 0°$ the experimental distribution still has a normalized signal of 0.45, while the theoretical distribution is at a normalized signal of 0.10. This difference at low scattering angles around the surface normal is visible for all scattering incident conditions. Due to the normalization of the distribution to a maximum of 1, this difference could be due to less scattering around the surface normal for the theoretical simulations or due to more specular scattering. In panel **(b)** the angular distribution for $E_{kin,i} = 1.92\,eV$ and $\Theta_i = 40°$ is shown. The maximum and the scattering at high polar angles compares favorably with the experiment, while there is a slightly lesser difference at small scattering polar angles compared to the scattering at $E_{\mathrm{kin,i}} = 0.99\,\mathrm{eV}$ and $\Theta_{\mathrm{i}} = 40°$. The scattering at $E_{\mathrm{kin,i}} = 0.99\,\mathrm{eV}$ and $\Theta_{\mathrm{i}} = 55°$ in panel **(c)** shows a similar difference for the maximum, while there is more scattering at higher angles compared to the experiment. For scattering below the specular angle, the experiment first shows a similar decline as for the scattering at $\Theta_{\mathrm{i}} = 40°$, which then ends in a plateau at around a normalized signal at 0.4. The theoretical distributions falls of more quickly and levels off to zero at around the surface normal at $\Theta_{\mathrm{s}} = 0°$. Scattering at $E_{\mathrm{kin,i}} = 1.92\,\mathrm{eV}$ and $\Theta_{\mathrm{i}} = 55°$ shows a similar behavior, both in experiment and theory as scattering at $E_{\mathrm{kin,i}} = 0.99\,\mathrm{eV}$, with the same differences between experiment and theory. For the maximum of the distribution and the scattering at higher angles than the specular angle show nearly no differences with the experiment.

Overall, the theoretical distributions compare well with the experiment at scattering angles around and higher than the specular angle, while there is a difference for scattering around the surface normal and at low scattering angles. The theoretical distributions fall off more quickly than the experimental ones going from the specular scattering toward the surface normal.

4.2.5 Discussion of the Differences between the Experiment and Theory

Overall there is a good agreement between experiment and theory, with all differences being relativly small adn the general shape of most distributions being well reproduced. The main differences between the experimental distributions and the theoretical ones are the consistent shift toward lower kinetic energies, the increased width of most distributions and the deviations for angular distributions close to scattering at the surface normal. During the prior analysis the quality of the HDNNP fit has been investigated and no major differences between the prediction of the HDNNP and the underlying DFT reference calculations have been found. Additionally, all of the differences between the experiment and theory have been present for the whole process of generating the HDNNP, in all generations where there were MD simulations using experimental conditions calculated. While it is not impossible that some of the differences are due to inaccuracies in the prediction of the HDNNP, it is highly improbable that the evident difference between the experimental and theoretical distributions is solely caused by disparities between the HDNNP PES and the RPBE PES. In the following section, the other possible sources for the differences will be analysed. There are four distinct sources from which the discrepancies between experiment and theory could stem from. 1) The HDNNP could be unable to correctly predict the RPBE PES, due to fitting errors or missing data. 2) The classical MD simulations could be unable to give the right description of the trajectory, due to non classical effects, like quantum effects or electronic excitation. 3) The RPBE functional could over or underestimate certain effects to lead to higher kinetic energy loss and different distributions. 4) And lastly the experiment could be adding unknown factors that change the distribution, like steps on the surface or adsorbed atoms and molecules. The influence of the HDNNP is the most accessible to test and were already discussed in the previous sections. Next the differences between the theoretical and experimental surface will be analyzed.

Discussion of the Influence of the Surface

As mentioned in section 2.2.4 the surface has been prepared in a manner where the termination should be well defined. There also should be no adsorbates on the surface, due to the ultra high vacuum conditions and annealing. Nonetheless, as already discussed, steps in the surface layers introduce a second termination which drastically changes how the theoretical distributions would look like. A high quantity of vacancies, ad-atoms or adsorbates could have a similar effect. A theoretical description of those changes of the structure of the surface is nearly impossible, due to multiple factors. First introducing multiple possible surface effects introduces more parameters for comparing the distributions. In the absence of knowledge regarding the specific nature and extent of the surface variations, it would be necessary to calculate distributions for all conceivable effects. Following this, the combination of distributions that most closely resembles the experimental data would need to be identified in order to determine the type and quantity of surface imperfections present. Additionally it would increase the amount of reference calculations drastically, adding additional calculations for the interaction between the H atom and every imperfection. This would also increase the amount of time necessary to validate that the HDNNP is accurate enough for every possible interaction. While the increase in cost, both in time and computing resources, could be managed, introducing more variables into the creation of the distribution would lead to multiple, equivalently accurate, solutions for the smallest difference between the experimental and theoretical distribution. Simply adding more variables on the computational side will not facilitate the identification of the sources of the observed deviations.

There are two differences between the experimental and theoretical kinetic energy distributions that could be caused by the discrepancies between the experimental and theoretical surface. The difference between the normalized signal at low scattering angles could be caused by a rougher experimental surface. In order for a H atom to scatter at an angle close to the surface normal it needs something to transfer its forward momentum to. In a classical collision model, the H atom would need to collide with a protrusion on the surface for it to scatter with a lower scattering angle, as seen in figure 4.26. The ideal [0001] surface, used for the calculations, should have the least possible amount of these protrusions. Steps, adatoms and vacancies should all increase the amount of scattering close to the surface normal. Increasing the scattering between the surface normal and the specular angle should reduce the differences between the theoretical and experimental angular distributions drastically. It could also slightly move the maximum of the distribution toward lower scattering angles, improving the agreement between experiment and theory even further. Another possible source of more scattering at angles around the surface normal is the desorption of adsorbed H atoms from the surface or from subsurface sites. During all simulation that have been run there was not a single incident of H atoms going below the surface. The first PBE data set included structures with H atoms below the surface and simulations of H atoms moving from sub surface positions, in principle, was possible. However, no such event was recorded. Analyzing the kinetic energy distribution for scattering at small scattering angles also does not show any signs of another form of scattering. If there were additional H atoms detected from desorption, they should have a different kinetic energy distribution compared to the directly scattered H atoms. Here, the agreement between the experimental distribution, which could include such events, and the theoretical distribution, which does not include such events, is the best from all scattering angels. If there were H atoms desorbing from the surface and being detected in the experiment, this should be evident when comparing the experimental and theoretical energy distributions.

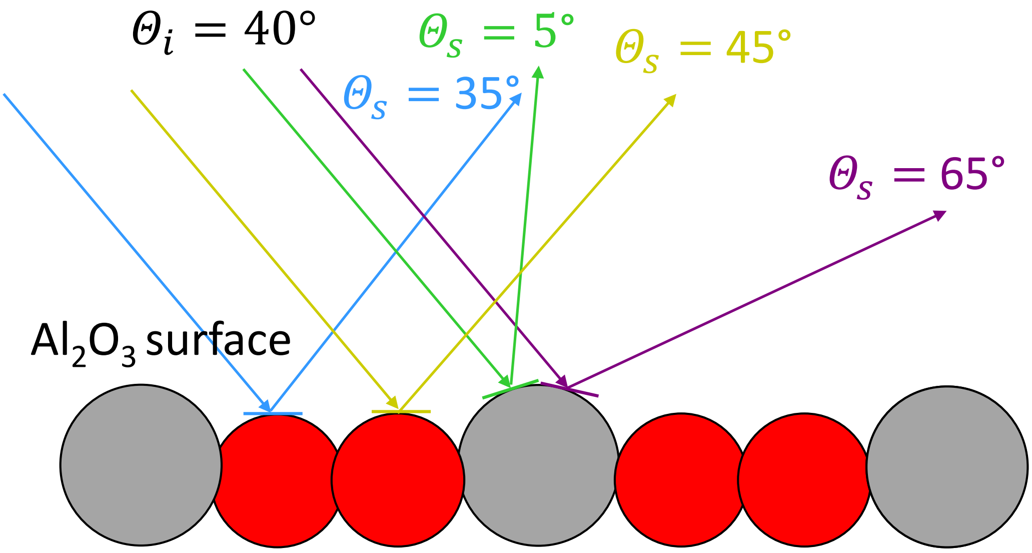


Fig. 4.26: Schematic representation on how the H atom could scatter from the surface under experimental scattering angles in a classical scattering model. All angles are correctly represented by the arrows with the same colours.

But the difference at small scattering angles could also be affected by a more subtle effects than desorption. When the H atom scatters from the surface it transfers some part of its kinetic energy toward the surface and due to the parallel momentum conversation, the scattering angle is affected by the normal momentum transfer. This means that energy transfer from the H atom to the surface will lead to scattering angles higher than the specular angle, while energy transfer from the surface to the H atom will increase the scattering at angles closer to the surface normal. Significant differences are observed between the experimental and theoretical distributions at high $E_{\mathrm{kin,s}}$, particularly in the case of scattering at $E_{\mathrm{kin,i}} = 0.99\,\mathrm{eV}$. Notably, the number of scattered H atoms with kinetic energies higher than $E_{\mathrm{kin,i}}$ is considerably greater in the experimental distribution compared to the theoretical one. If the theoretical surface is less effective at transferring kinetic energy toward the H atom, while also showing more transfer of kinetic energy from the H atom toward the surface, the theoretical angular distributions should be shifted toward higher scattering angles and show less scattering around the surface normal. This effect alone is likely not sufficient to explain the differences between the experimental and predicted angular distributions close to the surface normal, but it will definitely contribute toward it. This effect could be due to different surface structures having a varying efficacy at transferring kinetic energy toward the H atom, or it could come from differences in the surface motion, that would be caused by an inaccurate description of the surface motion by the functional or both. The surface can also have an influence on the width and the kinetic energy loss of the distributions. The kinetic energy loss is influenced by the effective mass of the surface during the scattering process, as discussed previously. Changing the surface structure by adding or removing atoms will influence the effective mass of the surface. If, for example, the experimental surface would have more Al atom compared to the ideal [0001] α-$\mathrm{Al_2O_3}$ surface, the mean kinetic energy loss should be smaller, since the mass of the Al atom is larger than the mass of the O atom. Here, analyzing from where the H atom is scattered from the surface could give insight regarding the plausibility of more Al atoms being present on the experimental surface. The simplified point of scattering can be plotted for the different kinetic energy distributions to identify the regions of the surface were the scattering originates from. Since there are many more scattered atoms per distribution compared to the previously plotted peak, the points need to be colored by the density of the points, giving a heat map of the point of scattering. This will show the regions from were the most H atoms are scattered at a glance. In figure 4.27, the heat map can be seen for $E_{\mathrm{kin,i}} = 1.92\,\mathrm{eV}$ and $\Theta_{\mathrm{i}} = 40°$ with all experimentally measured scattering angles.

The data is shown for one unit cell equivalent of the [0001] α-$\mathrm{Al_2O_3}$ surface, shown by the black lines. In the panel for $\Theta_{\mathrm{s}} = 5°$ most of the scattering is coming from scattering off the Al atom, which is in line with the classical scattering model, were a protrusion is needed to scatter at small scattering angles. A similar behavior can be seen for the scattering at $E_{\mathrm{kin,i}} = 0.99\,\mathrm{eV}$, $\Theta_{\mathrm{i}} = 40°$ and $\Theta_{\mathrm{s}} = 5°$ in figure 4.28, were most of the signal is again coming from scattering of the Al atom.

For scattering at those angles, the protrusion of the Al atom out of the surface is needed, whereas for the scattering at $\Theta_{\mathrm{s}} = 35°$ and 45°, close to the specular angle, the H atom can scatter from almost the whole surface. Except the position of the Al

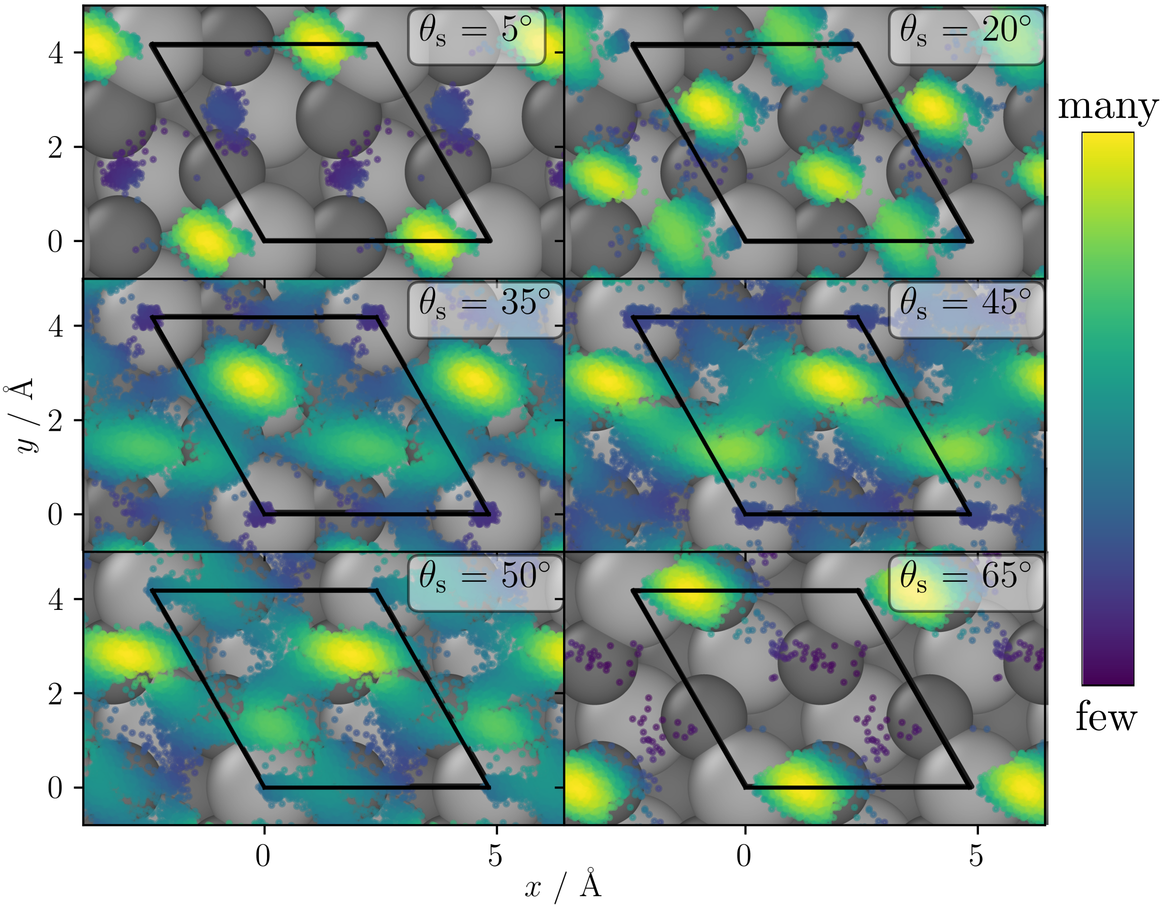


Fig. 4.27: Point of scattering heat map for the scattering at $E_{\mathrm{kin,i}} = 1.92\,\mathrm{eV}$ and $\Theta_\mathrm{i} = 40°$. The color of the points represents how many scattering events happen at that point. The point of scattering is determined by the lowest point in the trajectory. The grey scale background shows the [0001] α-Al_2O_3 surface and the black border frames in the unit cell.

atom, the surface seems to be flat in regards to the H atom scattering of it. The same can be seen for the scattering at $\Theta_\mathrm{i} = 55°$ and $\Theta_\mathrm{s} = 55°$, for both kinetic energies, in figure 4.29 and 4.30 were the scattering is happening from the whole surface.
And again for both kinetic energies the scattering at $\Theta_\mathrm{s} = 35°$, which needs a protrusion to stop the forward momentum, the H atom scatters of the Al atom. During the simulations it is evident that the surface favors scattering at the specular angle, while it is unfavorable for scattering at low scattering angles. This is in line with the observations for the angular distributions, where there is a strong mismatch between the amount of scattering at low scattering angles between the experiment and theory. The experimental surface had to be rougher than the theoretical surface, causing more scattering at low scattering angles, which cannot be seen for the theoretical surface. If the roughness is created by atoms that are heavier or more strongly bound than the O

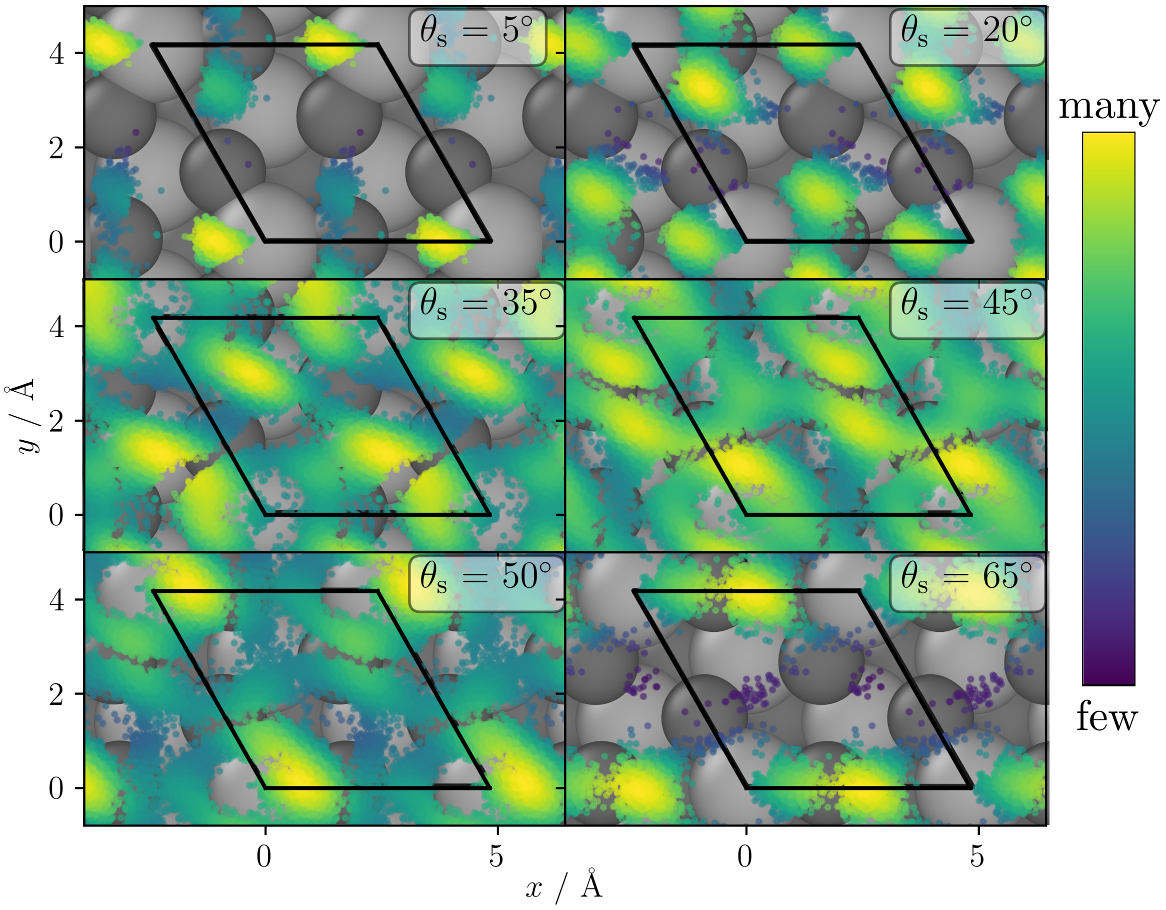


Fig. 4.28: Point of scattering heat map for the scattering at $E_{\mathrm{kin,i}} = 0.99\,\mathrm{eV}$ and $\Theta_\mathrm{i} = 40°$. The colour of the points represents how many scattering events happen at that point. The point of scattering is determined by the lowest point in the trajectory. The grey scale background shows the [0001] α-Al_2O_3 surface and the black border frames in the unit cell.

atom on the perfect [0001] α-Al_2O_3 surface, which most of the theoretical scattering is happening from, the smaller kinetic energy loss of the experimental distributions could also be explained. Reducing the space from which the H atom can scatter toward a certain scattering angle should also reduce the width of the distributions. For scattering at $E_{\mathrm{kin,i}} = 1.92\,\mathrm{eV}$ and $\Theta_\mathrm{i} = 40°$, the kinetic energy distributions which are closest in width compared to the experiment, scattering at $\Theta_\mathrm{s} = 5°$ and 65°, also show the smallest area from which the scattering can happen on the surface. Every site on the surface from which the scattering can happen should have a distribution of possible $E_{\mathrm{kin,s}}$, depending on the current positions of surface atoms and the correlating effective mass as well as scattering geometry. The distribution should look slightly different for each scattering site and the maximum should be slightly shifted depending on the environment of that site. The full kinetic energy distribution is then a composition of

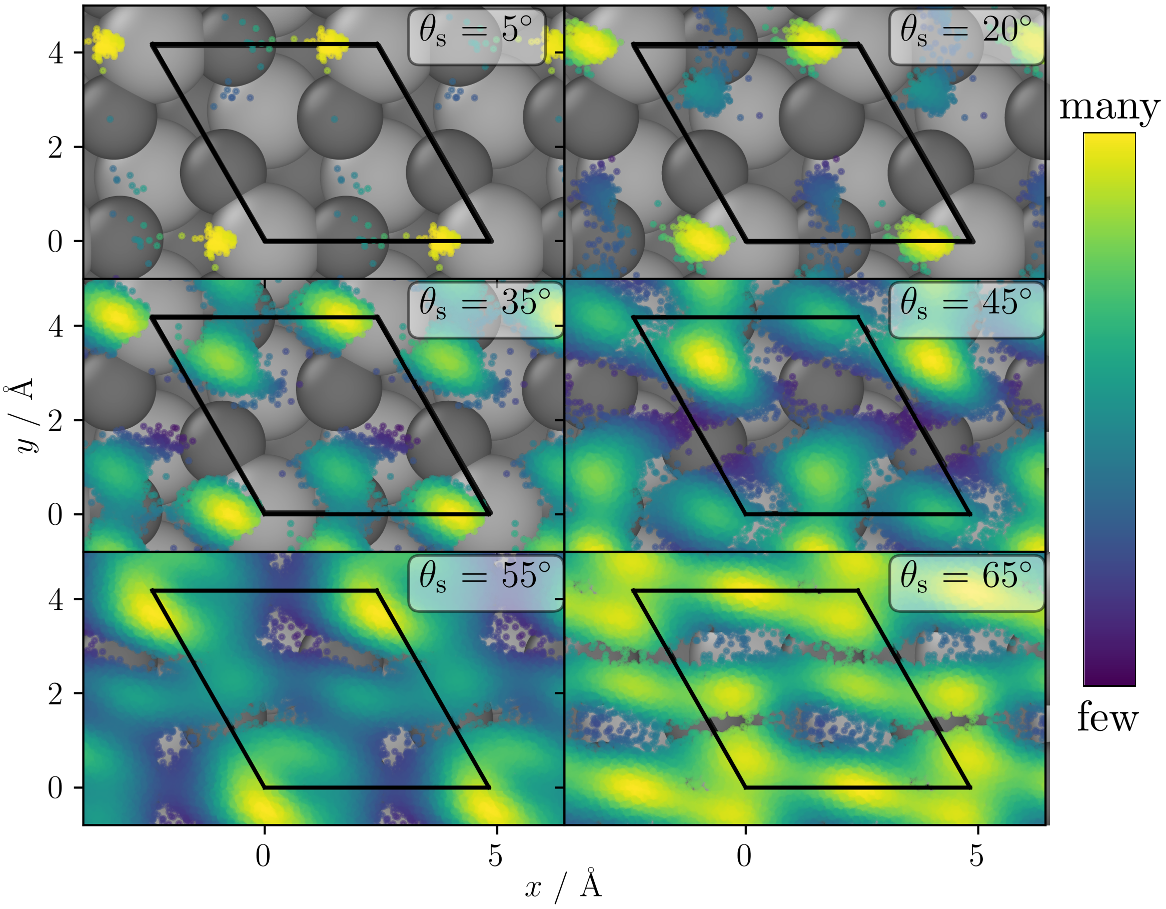


Fig. 4.29: Point of scattering heat map for the scattering at $E_{\mathrm{kin,i}} = 0.99\,\mathrm{eV}$ and $\Theta_{\mathrm{i}} = 55°$. The colour of the points represents how many scattering events happen at that point. The point of scattering is determined by the lowest point in the trajectory. The grey scale background shows the [0001] α-$\mathrm{Al_2O_3}$ surface and the black border frames in the unit cell.

all those distributions. Summing up the distributions of more different sites all leading to the same scattering angle should increase the width of the summed up kinetic energy distribution. A rougher surface would mean there is less possible space for scattering to happen at $\Theta_{\mathrm{s}} = 35°$ and $45°$, where the theoretical distribution is wider than the experimental one.

Overall, comparing the experimental and theoretical results shows that the scattering from the α-$\mathrm{Al_2O_3}$ surface is strongly affected by the exact geometries of the surface. The amount of scattering at certain angles is strongly correlated to the needed surface roughness, which can explain the difference in scattering signal for the angular distributions at low scattering polar angles. The different widths of the distributions could also be affected by the varying surface roughness, giving more sites for the scattering to happen at specular scattering angles, increasing the width of the experimental kinetic

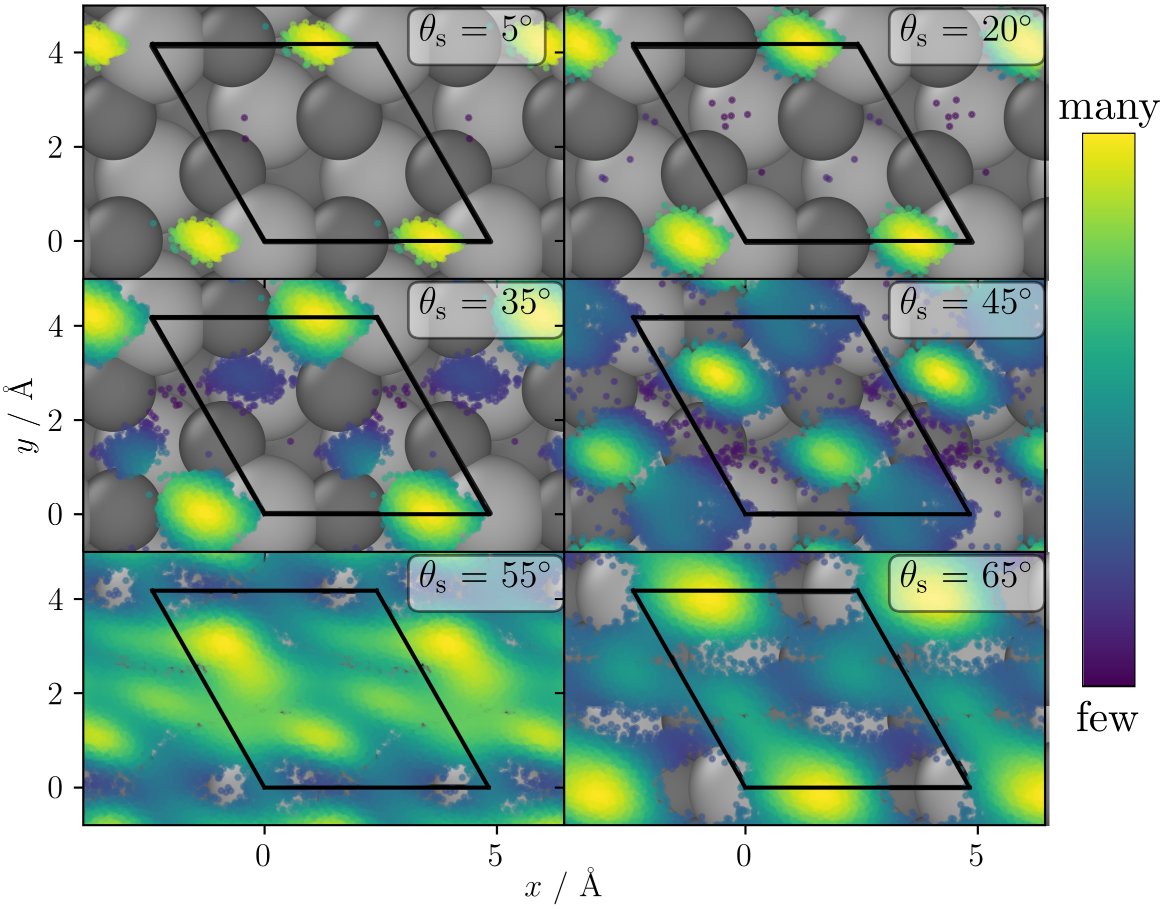


Fig. 4.30: Point of scattering heat map for the scattering at $E_{\mathrm{kin,i}} = 1.92\,\mathrm{eV}$ and $\Theta_\mathrm{i} = 55°$. The colour of the points represents how many scattering events happen at that point. The point of scattering is determined by the lowest point in the trajectory. The grey scale background shows the [0001] α-$\mathrm{Al_2O_3}$ surface and the black border frames in the unit cell.

energy distributions.

Discussion of the Influence of the Functional

Some of the differences between the experimental and theoretical distributions could be caused by the used functional with the energy loss being especially sensitive with regards to the PES. In the simulations the effective mass of the surface is determined by the strength of the bonds between the surface atoms and strength of the attractive and repulsive H atom interactions can also influence the amount of energy the H atom is transferring or receiving to and from the surface. Here a lot of different effects are working together to determine the kinetic energy loss during the simulation. The main observation is that the kinetic energy loss during the simulation is higher than experimentally measured one. This could be either caused by a too strong interaction between the H atom and the surface or too softly bound surface atoms. Here, not only the surface in general could be too softly bound, but also the interaction between the H atom and the Al atoms or the O atoms could be separately lead to too much kinetic energy loss.

The best agreement between the experimental and theoretical kinetic energy distributions is obtained at low scattering angles with high kinetic energies. The biggest difference in mean kinetic energy is loss present at high scattering angles and low kinetic incident energies, especially the scattering at $E_{\mathrm{kin,i}} = 0.99\,\mathrm{eV}$ and $\Theta_{\mathrm{s}} = 65°$. Both cases represent opposite extremes of H atom interaction with a surface. When the H atom scatters at high kinetic energies and low scattering angles, the simulation time and interaction time inside the surface potential are minimal. However, when the scattering angle is high and kinetic energy is low, the simulation and interaction time with the surface are the longest compared to all other simulations. The longer interaction time could increase the effect of the stronger interaction, explaining why the kinetic energy distributions for high kinetic energies are in better agreement with the experiment compared to the distributions for low kinetic energies. It would also explain why the distributions at low kinetic incident conditions are all more similar in shape to each other. If a big part of the kinetic energy loss is not happening during the collision, but when the H atom is inside of the potential of the surface, the influence of the surface structure and the scattering conditions on the shape of the distribution is reduced, resulting in similar shaped distributions for all scattering angles. It would also explain why the maxima of the distributions for both the scattering at $E_{\mathrm{kin,i}} = 0.99\,\mathrm{eV}$, $\Theta_{\mathrm{i}} = 40°$ and $E_{\mathrm{kin,i}} = 0.99\,\mathrm{eV}$, $\Theta_{\mathrm{i}} = 55°$ are not increasing with the scattering angle, which the binary collision model would predict. Moreover, for the scattering at $E_{\mathrm{kin,i}} = 1.92\,\mathrm{eV}$, the influence of the attractive potential is reduced and the prediction of the binary collision model fits better.

The higher kinetic energy loss could also be caused by the softer description of α-Al_2O_3 by the functional. Comparing the bulk modulus calculated with the RPBE functional with the experimentally measured one is a difference of 10%, as can be seen in table 4.1. It is unknown how well this difference in bulk modulus is transferred to the local hardness of the surface during the scattering event. Nonetheless, the surface hardness could be underestimated by the RPBE functional, which would lead to more kinetic energy loss.

The plausibility of an over or underestimation of the interaction between the surface and the H atom can be checked by comparing the interaction of the H atom with the surface calculated using different functionals. In figure 4.31 the 1D energy curve for the H atom adsorbing onto the topmost Al atom is shown for three different functionals. For each functional, the 2×2 super cell was created using the same approach as described for the RPBE structure in section 3.3.1. The bulk unit cell was optimized using the respective functional. Then the slab structure was created from the bulk structure.

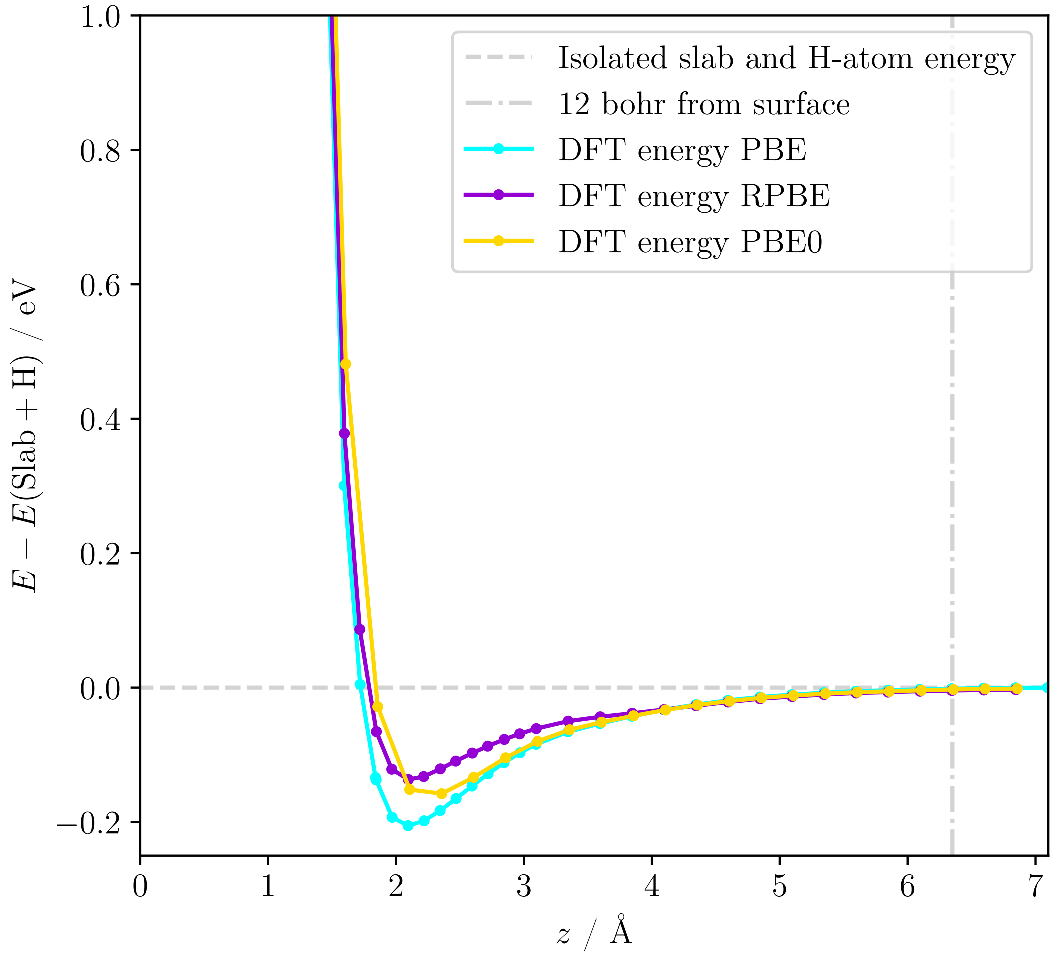


Fig. 4.31: 1D energy curve for the H atom directly on top of the Al atom of a optimized α-Al_2O_3 slab. The adsorption energy is shown for the PBE, RPBE and PBE0 functional and plotted against the distance from the surface z.

All three functionals show different adsorption energies and minima locations. The RPBE functional has the smallest adsorption energy of 0.11 eV, while the PBE functional has the largest adsorption energy with 0.20 eV. The PBE0 functional has a adsorption energy of 0.14 eV and the position of the minimum is shifted from 2.1 Å for the RPBE and PBE functional toward 2.3 Å from the surface for the PBE0 functional. The figure is showing the adsorption onto the optimized surface for all three functionals. The difference in adsorption energy is similar in magnitude to the difference in mean kinetic energy between experiment and theory. It is not unlikely that the interaction of the H atom with the surface is not accurately described by the RPBE functional. In order to achieve better agreement between the experimental and predicted energy distributions other functionals could be employed in the future.

5 Conclusion

In this thesis a HDNNP describing the scattering of H atoms from the [0001] α-Al_2O_3 surface was created. It is based on RPBE DFT reference calculations. After careful validation it is applied to further the understanding of both the theoretical side of constructing and using MLP in combination with experimental data, as well as deepening the understanding of the conditions under which the experiment was measured.

The careful construction of the reference data set, starting with the structural setup of the bulk and slab calculations, is described in the computational details part of this thesis. The discovery of the shift of H atom spin towards the opposite side of the slab demanded a reevaluation of the structural setup. This was guided by using the fitting capabilities of the HDNNP to detect inconsistencies in the data set. The investigation was supported by the use of the ACSFs to determine if the HDNNP is able to differentiate the H atom spin states resulting from different chemical environments. After reducing the amount of H atoms per slab to one and freezing the bottom four Al-O-O layers of the slab, the RPBE H atom spin states were accurately differentiated by the HDNNP.

The grid approach for the initial structure generation of H atoms interacting with the slab proved to be useful in quickly and efficiently generating a potential which is able to run the first MD simulations for starting the active learning process. Furthermore it enabled the visualization of the PES, helping in the validation of the potential. The structures for the reference calculations were mostly created using active learning. By running MD simulations with preliminary HDNNPs, missing structures were found and holes in the configurational space were filled without the need of expensive ab-initio MD calculations. The availability of the preliminary HDNNPs also allowed to discover problems with the data set on the fly, reducing their impact by discovering them early. The change in structural setup due to the H atom spin problems, then was accelerated by using the HDNNP fitted to the old reference data set and furthest point sampling to quickly generate the new structures.

The the HDNNP was validated by comparing multiple different feature of the system with DFT reference data. The phonon band structure comparison showed that the HDNNP is able to reproduce the possible energy transfer mechanism of the system responsible for transferring the translation energy of the H atom onto the surface and then distributing the energy over the structure. The visualization of the PES using the aforementioned H atom grid structures showed a uniformly accurate coverage of the energy prediction of the HDNNP over the whole unit cell, while the 1D-energy plots showed that the HDNNP is able to reproduce the RPBE energies and forces for energy ranges occurring during the MD simulations. Using selected ab-initio MD simulations it could be shown that the HDNNP is able to reproduce the energies and forces for scattering trajectories, differing in high and low scattering kinetic energies and angles close to the surface normal or the specular angle. While the amount of ab-inito MD simulations is too small for a statistical evaluation, they are nonetheless a important part of the validation process.

While the validation using the RPBE calculations shows how well the the HDNNP is able to reproduce reference calculations, the goal is to describe the system under real conditions. To benchmark the ability of the HDNNP to predict experimental results, experimental kinetic energy distributions and angular distributions for four sets of incident conditions were compared to the results of MD simulations using the HDNNP. By comparing theoretical kinetic energy distributions at $\phi_i = 0°$ and $180°$ with the experimental kinetic energy distributions it could be determined that the scattering in the experiment is a composition of two oxygen layerings, which are energetically equivalent and interchangeable by a mirror operation. Overall the theoretical results agree with the experiment, only showing one additional peak at $E_{\mathrm{kin,i}} = 1.92\,\mathrm{eV}$, $\Theta_{\mathrm{i}} = 40°$ and $\Theta_{\mathrm{s}} = 65°$, which is most likely a feature of the RPBE PES. The results of the HDNNP tend to overestimate the kinetic energy loss by around $0.1\,\mathrm{eV}$ and show broader distributions for the specular scattering. In combination with the big difference in H atom flux for scattering angels around the surface normal, it has to be expected that the differences stem from the difference in surface roughness. A rougher surface leads to more scattering around the surface normal and reduces the possible geometries at which scattering at the specular angle is possible. The higher kinetic energy loss is most likely caused by the overestimation of the attractive interaction between the H atom and the surface. Especially for the scattering at $E_{\mathrm{kin,i}} = 0.99\,\mathrm{eV}$ the resulting kinetic energy loss seems to be less determined by the scattering geometry and more by attractive potential.

Overall the findings demonstrate the ability of the HDNNP to reproduce the reference calculations and the associated PES for solid-gas interfaces under vacuum conditions and how they can be used to broaden the availability of systems which can be investigated theoretically. It shows where there is still need to investigate electronic structure methods and how important well defined experimental conditions are for benchmarking theory.

6 Outlook

The HDNNP developed in this thesis can be applied to further investigate the interaction between a H atom and the [0001] α-Al_2O_3 surface. Without further adjustment of the potential, investigations of the sticking of H atoms, as well as the scattering under different incident conditions are possible. Increasing the initial kinetic energy of the H atom improves the comparison between the experiment and theory for scattering at $E_{\mathrm{kin,i}} = 1.92\,\mathrm{eV}$ but only slightly improves the description of the scattering at $E_{\mathrm{kin,i}} = 0.99\,\mathrm{eV}$. Since the experimental spectra shows a significant contribution of H atoms with $E_{\mathrm{kin,s}} >= E_{\mathrm{kin,i}}$ it would be interesting to see if increasing the surface temperature could improve the comparison between the theoretical and experimental results. Since the HDNNP is able to describe the H atom adsorbed on the surface, and investigations of H atom mobility on the surface could be done using the potential.

The HDNNP could be further improved by changing the functional used for the reference calculations to a hybrid functional, increasing the accuracy of the reference calculations. Test calculations using the PBE0 functional, shown in figure 4.31 and table 6.1, show that the prediction of the bulk modulus by the PBE0 functional is significantly more accurate compared to the experimental value. The increased bulk modulus could lead to less kinetic energy loss of the H atom during the simulations. The position of the adsorption minimum is also slightly different between the PBE and RPBE functional, hinting towards a slightly different description of the H atom interaction with the surface, which could lead to a improved description of the experimental data.

The configurational space of the HDNNP could also be extended to either include steps or point defects to more closely resemble the experimental surface and investigate how those surface defects can influence the scattering. Increasing the diversity of the available experimental measurements could also greatly help in determining the sources of the deviations between the experiment and theory.

Tab. 6.1: The calculated and experimental bulk modulus B_0 in GPa.

Parameter	RPBE	PBE	PBE0	Exp.
B_0	222	229	258	253

References

[1] Z. Ma, F. Zaera in *Encyclopedia of Inorganic Chemistry*, John Wiley & Sons, Ltd, **2006**.

[2] J. G. Chen, R. M. Crooks, L. C. Seefeldt, K. L. Bren, R. M. Bullock, M. Y. Darensbourg, P. L. Holland, B. Hoffman, M. J. Janik, A. K. Jones, M. G. Kanatzidis, P. King, K. M. Lancaster, S. V. Lymar, P. Pfromm, W. F. Schneider, R. R. Schrock, Beyond fossil fuel-driven nitrogen transformations, *Science* **2018**, *360*, eaar6611.

[3] I. Fechete, Y. Wang, J. C. Védrine, The past, present and future of heterogeneous catalysis, *Catalysis Today* **2012**, *189*, Catalytic Materials for Energy: Past, Present and Future, 2–27.

[4] L. Grajciar, C. J. Heard, A. A. Bondarenko, M. V. Polynski, J. Meeprasert, E. A. Pidko, P. Nachtigall, Towards operando computational modeling in heterogeneous catalysis, *Chem. Soc. Rev.* **2018**, *47*, 8307–8348.

[5] C. M. Handley, P. L. A. Popelier, Potential Energy Surfaces Fitted by Artificial Neural Networks, *The Journal of Physical Chemistry A* **2010**, *114*, PMID: 20131763, 3371–3383.

[6] J. Behler, Perspective: Machine learning potentials for atomistic simulations, *J. Chem. Phys.* **2016**, *145*, 170901.

[7] V. Botu, R. Batra, J. Chapman, R. Ramprasad, Machine Learning Force Fields: Construction, Validation, and Outlook, *The Journal of Physical Chemistry C* **2017**, *121*, 511–522.

[8] D. Shanavas Rasheeda, A. Martín Santa Daría, B. Schröder, E. Mátyus, J. Behler, High-dimensional neural network potentials for accurate vibrational frequencies: the formic acid dimer benchmark, *Phys. Chem. Chem. Phys.* **2022**, *24*, 29381–29392.

[9] M. Eckhoff, F. Schönewald, M. Risch, C. A. Volkert, P. E. Blöchl, J. Behler, Closing the gap between theory and experiment for lithium manganese oxide spinels using a high-dimensional neural network potential, *Phys. Rev. B* **2020**, *102*, 174102.

[10] S. Wille, H. Jiang, O. Bünermann, A. M. Wodtke, J. Behler, A. Kandratsenka, An experimentally validated neural-network potential energy surface for H-atom on free-standing graphene in full dimensionality, *Phys. Chem. Chem. Phys.* **2020**, *22*, 26113–26120.

[11] O. Bünermann, H. Jiang, Y. Dorenkamp, A. Kandratsenka, S. M. Janke, D. J. Auerbach, A. M. Wodtke, Electron-hole pair excitation determines the mechanism of hydrogen atom adsorption, *Science* **2015**, *350*, 1346–1349.

[12] L. Lecroart, N. Hertl, Y. Dorenkamp, H. Jiang, T. N. Kitsopoulos, A. Kandratsenka, O. Bünermann, A. M. Wodtke, Adsorbate modification of electronic nonadiabaticity: H atom scattering from p(2 × 2) O on Pt(111), *The Journal of Chemical Physics* **2021**, *155*, 034702, DOI `10.1063/5.0058789`.

[13] N. Hertl, A. Kandratsenka, O. Bünermann, A. M. Wodtke, Multibounce and Subsurface Scattering of H Atoms Colliding with a van der Waals Solid, *The Journal of Physical Chemistry A* **2021**, *125*, PMID: 34181858, 5745–5752.

[14] M. Andersson, F. Abild-Pedersen, I. Remediakis, T. Bligaard, G. Jones, J. Engbæk, O. Lytken, S. Horch, J. Nielsen, J. Sehested, J. Rostrup-Nielsen, J. Nørskov, I. Chorkendorff, Structure sensitivity of the methanation reaction: H2-induced CO dissociation on nickel surfaces, *Journal of Catalysis* **2008**, *255*, 6–19.

[15] D. Hollenbach, E. E. Salpeter, Surface Adsorption of Light Gas Atoms, *The Journal of Chemical Physics* **2003**, *53*, 79–86.

[16] Z. Medina, B. Jackson, Quantum studies of light particle trapping, sticking, and desorption on metal and graphite surfaces, *The Journal of Chemical Physics* **2008**, *128*, 114704, DOI `10.1063/1.2890043`.

[17] G.-J. Kroes, M. Pavanello, M. Blanco-Rey, M. Alducin, D. J. Auerbach, Ab initio molecular dynamics calculations on scattering of hyperthermal H atoms from Cu(111) and Au(111), *The Journal of Chemical Physics* **2014**, *141*, 054705, DOI `10.1063/1.4891483`.

[18] S. Kaufmann, D. Schwarzer, C. Reichardt, A. M. Wodtke, O. Bünermann, Generation of ultra-short hydrogen atom pulses by bunch-compression photolysis, *Nature Communications* **2014**, *5*, 5373.

[19] S.-H. Cho, N. D. Walther, S. T. Nguyen, J. T. Hupp, Anodic aluminium oxide catalytic membranes for asymmetric epoxidation, *Chem. Commun.* **2005**, 5331–5333.

[20] T. Riittonen, E. Toukoniitty, D. K. Madnani, A.-R. Leino, K. Kordas, M. Szabo, A. Sapi, K. Arve, J. Wärnå, J.-P. Mikkola, One-Pot Liquid-Phase Catalytic Conversion of Ethanol to 1-Butanol over Aluminium Oxide—The Effect of the Active Metal on the Selectivity, *Catalysts* **2012**, *2*, 68–84.

[21] J. Behler, M. Parrinello, Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces, *Phys. Rev. Lett.* **2007**, *98*, 146401.

[22] R. S. F. Peter W. Atkins, *Molecular Quantum Mechanics*, Oxford University Press, **2010**.

[23] M. Born, Zur Quantenmechanik der Stoßvorgänge, *Z. Phys.* **1926**, *37*, 863–867.

[24] W. Pauli, Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren, *Z. Phys.* **1925**, *31*, 765–783.

[25] W. Pauli, The Connection Between Spin and Statistics, *Phys. Rev.* **1940**, *58*, 716–722.

[26] E. Schrödinger, Quantisierung als Eigenwertproblem, *Ann. Phys.* **1926**, *384*, 361–376.

[27] E. Schrödinger, Quantisierung als Eigenwertproblem, *Ann. Phys.* **1926**, *384*, 489–527.

[28] M. Born, R. Oppenheimer, Zur Quantentheorie der Molekeln, *Ann. Phys.* **1927**, *389*, 457–484.

[29] P. Hohenberg, W. Kohn, Inhomogeneous Electron Gas, *Phys. Rev. B* **1964**, *136*, 864–871.

[30] L. H. Thomas, The calculation of atomic fields, *Math. Proc. Camb. Philos. Soc.* **1927**, *23*, 542–548.

[31] E. Fermi, Eine statistische Methode zur Bestimmung einiger Eigenschaften des Atoms und ihre Anwendung auf die Theorie des periodischen Systems der Elemente, *Z. Phys.* **1928**, *48*, 73–79.

[32] D. R. Hartree, The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part I. Theory and Methods, *Math. Proc. Camb. Philos. Soc.* **1928**, *24*, 89–110.

[33] V. Fock, Näherungsmethode zur Lösung des quantenmechanischen Mehrkörperproblems, *Z. Phys.* **1930**, *61*, 126–148.

[34] J. P. Perdew, K. Schmidt, Jacob's ladder of density functional approximations for the exchange-correlation energy, *AIP Conf. Proc.* **2001**, *577*, 1–20.

[35] C. R. Jacob, M. Reiher, Spin in Density-Functional Theory, *Int. J. Quantum Chem.* **2012**, *112*, 3661–3684.

[36] J. P. Perdew, K. Burke, M. Ernzerhof, Generalized Gradient Approximation Made Simple, *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

[37] B. Hammer, L. B. Hansen, J. K. Nørskov, Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals, *Phys. Rev. B* **1999**, *59*, 7413–7421.

[38] C. Adamo, V. Barone, Toward reliable density functional methods without adjustable parameters: The PBE0 model, *J. Chem. Phys.* **1999**, *110*, 6158–6170.

[39] F. Bloch, Über die Quantenmechanik der Elektronen in Kristallgittern, *Z. Phys.* **1929**, *52*, 555–600.

[40] L. Brillouin, E. Rabinowitsch, *Die Quantenstatistik und Ihre Anwendung auf die Elektronentheorie der Metalle*, Springer, Berlin, **1931**.

[41] E. Wigner, F. Seitz, On the Constitution of Metallic Sodium, *Phys. Rev.* **1933**, *43*, 804–810.

[42] Q. Wu, W. Yang, Empirical correction to density functional theory for van der Waals interactions, *J. Chem. Phys.* **2002**, *116*, 515–524.

[43] S. Grimme, Accurate description of van der Waals complexes by density functional theory including empirical corrections, *J. Comput. Chem.* **2004**, *25*, 1463–1473.

[44] A. Tkatchenko, M. Scheffler, Accurate Molecular Van Der Waals Interactions from Ground-State Electron Density and Free-Atom Reference Data, *Phys. Rev. Lett.* **2009**, *102*, 073005.

[45] F. L. Hirshfeld, Bonded-Atom Fragments for Describing Molecular Charge Densities, *Theoret. Chim. Acta* **1977**, *44*, 129–138.

[46] A. P. Bartók, S. De, C. Poelking, N. Bernstein, J. R. Kermode, G. Csányi, M. Ceriotti, Machine learning unifies the modeling of materials and molecules, *Sci. Adv.* **2017**, *3*, e1701816.

[47] V. L. Deringer, M. A. Caro, G. Csányi, Machine Learning Interatomic Potentials as Emerging Tools for Materials Science, *Adv. Mater.* **2019**, *31*, 1902765.

[48] F. Noé, A. Tkatchenko, K.-R. Müller, C. Clementi, Machine Learning for Molecular Simulation, *Ann. Rev. Phys. Chem.* **2020**, *71*, 361–390.

[49] J. Behler, G. Csányi, Machine Learning Potentials for Extended Systems - A Perspective, *Eur. Phys. J. B* **2021**, *94*, 142.

[50] J. Behler, Four Generations of High-Dimensional Neural Network Potentials, *Chem. Rev.* **2021**, *121*, 10037–10072.

[51] T. B. Blank, S. D. Brown, A. W. Calhoun, D. J. Doren, Neural network models of potential energy surfaces, *J. Chem. Phys.* **1995**, *103*, 4129–4137.

[52] C. M. Handley, P. L. A. Popelier, Potential Energy Surfaces Fitted by Artificial Neural Networks, *J. Phys. Chem. A* **2010**, *114*, 3371–3383.

[53] J. Behler, Neural network potential-energy surfaces in chemistry: a tool for large-scale simulations, *Phys. Chem. Chem. Phys.* **2011**, *13*, 17930–17955.

[54] J. Behler, Representing potential energy surfaces by high-dimensional neural network potentials, *J. Phys.: Condens. Matter* **2014**, *26*, 183001.

[55] J. Behler, Constructing High-Dimensional Neural Network Potentials: A Tutorial Review, *Int. J. Quantum Chem.* **2015**, *115*, 1032–1050.

[56] J. Behler, First Principles Neural Network Potentials for Reactive Simulations of Large Molecular and Condensed Systems, *Angew. Chem. Int. Ed.* **2017**, *56*, 12828–12840.

[57] A. P. Bartók, M. C. Payne, R. Kondor, G. Csányi, Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons, *Phys. Rev. Lett.* **2010**, *104*, 136403.

[58] A. V. Shapeev, Moment Tensor Potentials: a class of systematically improvable interatomic potentials, *Multiscale Model. Simul.* **2016**, *14*, 1153–1173.

[59] R. M. Balabin, E. I. Lomakina, Support vector machine regression (LS-SVM)—an alternative to artificial neural networks (ANNs) for the analysis of quantum chemistry data?, *Phys. Chem. Chem. Phys.* **2011**, *13*, 11710–11718.

[60] A. P. Thompson, L. P. Swiler, C. R. Trott, S. M. Foiles, G. J. Tucker, Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials, *J. Comp. Phys.* **2015**, *285*, 316–330.

[61] R. Drautz, Atomic cluster expansion for accurate and transferable interatomic potentials, *Phys. Rev. B* **2019**, *99*, 014104.

[62] J. Behler, Atom-centered symmetry functions for constructing high-dimensional neural network potentials, *J. Chem. Phys.* **2011**, *134*, 074106.

[63] M. Rupp, A. Tkatchenko, K.-R. Müller, O. A. von Lilienfeld, Fast and Accurate Modeling of Molecular Atomization Energies with Machine Learning, *Phys. Rev. Lett.* **2012**, *108*, 058301.

[64] L. Zhu, M. Amsler, T. Fuhrer, B. Schaefer, S. Faraji, S. Rostami, S. A. Ghasemi, A. Sadeghi, M. Grauzinyte, C. Wolverton, S. Goedecker, A fingerprint based metric for measuring similarities of crystalline structures, *The Journal of Chemical Physics* **2016**, *144*, 034203, DOI `10.1063/1.4940026`.

[65] A. P. Bartók, R. Kondor, G. Csányi, On representing chemical environments, *Phys. Rev. B* **2013**, *87*, 184115.

[66] M. F. Langer, A. Goessmann, M. Rupp, Representations of molecules and materials for interpolation of quantum-mechanical simulations via machine learning, *arXiv:2003.12081 [physics.comp-ph]* **2020**.

[67] Z. Deng, C. Chen, X.-G. Li, S. P. Ong, An electrostatic spectral neighbor analysis potential for lithium nitride, *npj Comput. Mater.* **2019**, *5*, 75.

[68] N. Artrith, T. Morawietz, J. Behler, High-dimensional neural-network potentials for multicomponent systems: Applications to zinc oxide, *Phys. Rev. B* **2011**, *83*, 153101.

[69] T. Morawietz, V. Sharma, J. Behler, A neural network potential-energy surface for the water dimer based on environment-dependent atomic energies and charges, *J. Chem. Phys.* **2012**, *136*, 064103.

[70] K. Yao, J. E. Herr, D. W. Toth, R. Mckintyre, J. Parkhill, The TensorMol-0.1 model chemistry: a neural network augmented with long-range physics, *Chem. Sci.* **2018**, *9*, 2261–2269.

[71] S. A. Ghasemi, A. Hofstetter, S. Saha, S. Goedecker, Interatomic potentials for ionic systems with density functional accuracy based on charge densities obtained by a neural network, *Phys. Rev. B* **2015**, *92*, 045131.

[72] X. Xie, K. A. Persson, D. W. Small, Incorporating Electronic Information into Machine Learning Potential Energy Surfaces via Approaching the Ground-State Electronic Energy as a Function of Atom-Based Electronic Populations, *J. Chem. Theory Comput.* **2020**, *16*, 4256–4270.

[73] T. W. Ko, J. A. Finkler, S. Goedecker, J. Behler, A fourth-generation high-dimensional neural network potential with accurate electrostatics including non-local charge transfer, *Nat. Commun.* **2021**, *12*, 398.

[74] T. W. Ko, J. A. Finkler, S. Goedecker, J. Behler, General-Purpose Machine Learning Potentials Capturing Nonlocal Charge Transfer, *Acc. Chem. Res.* **2021**, *54*, 808–817.

[75] M. Liebetrau, Master's thesis, Georg-August Universität Göttingen, **2018**.

[76] R. E. Kalman, A New Approach to Linear Filtering and Prediction Problems, *J. Basic Eng.* **1960**, *82*, 35–45.

[77] T. B. Blank, S. D. Brown, Adaptive, Global, Extended Kalman Filters For Training Feedforward Neural Networks, *J. Chemom.* **1994**, *8*, 391–407.

[78] E. Kocer, T. W. Ko, J. Behler, Neural Network Potentials: A Concise Overview of Methods, *Annual Review of Physical Chemistry* **2022**, *73*, PMID: 34982580, 163–186.

[79] B. J. Alder, T. E. Wainwright, Phase Transition for a Hard Sphere System, *J. Chem. Phys.* **1957**, *27*, 1208–1209.

[80] B. J. Alder, T. E. Wainwright, Studies in Molecular Dynamics. I. General Method, *J. Chem. Phys.* **1959**, *31*, 459–466.

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

[82] S. Nosé, A molecular dynamics method for simulations in the canonical ensemble, *Mol. Phys.* **1984**, *52*, 255–268.

[83] S. Nosé, A unified formulation of the constant temperature molecular dynamics methods, *J. Chem. Phys.* **1984**, *81*, 511–519.

[84] W. G. Hoover, Canonical dynamics: Equilibrium phase-space distributions, *Phys. Rev. A* **1985**, *31*, 1695–1697.

[85] G. J. Martyna, M. L. Klein, M. Tuckerman, Nosé-Hoover chains: The canonical ensemble via continuous dynamics, *J. Chem. Phys.* **1992**, *4*, 2635–2643.

[86] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, M. Scheffler, Ab initio molecular simulations with numeric atom-centered orbitals, *Comput. Phys. Commun.* **2009**, *180*, 2175–2196.

[87] FHI-aims – Fritz-Haber-Institute ab initio molecular simulations package, `https://aimsclub.fhi-berlin.mpg.de`, (October 23, 2016).

[88] J. P. Perdew, M. Ernzerhof, K. Burke, Rationale for mixing exact exchange with density functional approximations, *J. Chem. Phys.* **1996**, *105*, 9982–9985.

[89] C. G. Broyden, The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations, *J. Inst. Math. Appl.* **1970**, *6*, 76–90.

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

[91] D. Goldfarb, A Family of Variable-Metric Methods Derived by Variational Means, *Math. Comp.* **1970**, *24*, 23–26.

[92] D. F. Shanno, Conditioning of Quasi-Newton Methods for Function Minimization, *Math. Comp.* **1970**, *24*, 647–656.

[93] S. Plimpton, Fast Parallel Algorithms for Short-Range Molecular Dynamics, *J. Comput. Phys.* **1995**, *117*, 1–19.

[94] T. Morawietz, A. Singraber, C. Dellago, J. Behler, How van der Waals interactions determine the unique properties of water, *Proc. Natl. Acad. Sci.* **2016**, *113*, 8368–8373.

[95] A. Singraber, n2p2 – A neural network potential package, `https://github.com/CompPhysVienna/n2p2`, (December 9, 2019).

[96] A. Togo, I. Tanaka, First principles phonon calculations in materials science, *Scripta Materialia* **2015**, *108*, 1–5.

[97] A. Togo, First-principles Phonon Calculations with Phonopy and Phono3py, *Journal of the Physical Society of Japan* **2023**, *92*, 012001.

[98] D. P. Woodruff, Quantitative Structural Studies Of Corundum and Rocksalt Oxide Surfaces, *Chemical Reviews* **2013**, *113*, 3863–3886.

[99] Y. Eldar, M. Lindenbaum, M. Porat, Y. Zeevi, The farthest point strategy for progressive image sampling, *IEEE Transactions on Image Processing* **1997**, *6*, 1305–1315.

[100] J. Behler, RuNNer, `http://gitlab.com/TheochemGoettingen/RuNNer`, (August 22, 2019).

[101] F. D. Murnaghan, The Compressibility of Media under Extreme Pressures, *Proceedings of the National Academy of Sciences* **1944**, *30*, 244–247.

[102] F. Birch, Finite Elastic Strain of Cubic Crystals, *Phys. Rev.* **1947**, *71*, 809–824.

[103] X.-L. Wang, C. R. Hubbard, K. B. Alexander, P. F. Becher, J. A. Fernandez-Baca, S. Spooner, Neutron Diffraction Measurements of the Residual Stresses in Al2O3-ZrO2 (CeO2) Ceramic Composites, *Journal of the American Ceramic Society*, *77*, 1569–1575.

[104] D. Kulginov, M. Persson, C. T. Rettner, D. S. Bethune, An Empirical Interaction Potential for the Ar/Pt(111) System, *The Journal of Physical Chemistry* **1996**, *100*, 7919–7927.

[105] J. C. Tully, Washboard model of gas–surface scattering, *The Journal of Chemical Physics* **1990**, *92*, 680–686.

Appendix

A.1 RuNNer Settings

Tab. A.1: Values of the radial ACSF parameter η.

Element pair	$\eta\,/\,a_0^{-2}$
H-O	0, 0.005, 0.013, 0.029, 0.067, 0.187
O-H	0, 0.005, 0.013, 0.029, 0.067, 0.187
H-Al	0, 0.005, 0.012, 0.027, 0.058, 0.145
Al-H	0, 0.005, 0.012, 0.027, 0.058, 0.145
O-O	0, 0.005, 0.011, 0.023, 0.048, 0.109
O-Al	0, 0.004, 0.010, 0.019, 0.036, 0.070
Al-O	0, 0.004, 0.010, 0.019, 0.036, 0.070
Al-Al	0, 0.003, 0.008, 0.014, 0.024, 0.041

Tab. A.2: List of missing ACSFs from all combinations of the parameters $\lambda = \{-1, 1\}$, $\zeta = 1, 2, 4, 16$, and $\eta = 0\ a_0^{-2}, 0.05\ a_0^{-2}$.

Element triple	$\eta\,/\,a_0^{-2}$	λ	ζ
H-Al-Al	0.05	1	16
H-Al-Al	0.05	-1	16
Al-Al-Al	0.05	1	16
Al-Al-H	0.05	1	16
Al-Al-Al	0.05	-1	16
Al-Al-H	0.05	-1	16

Tab. A.3: RuNNer HDNNP construction settings excluding the specification of symfunction_short.

	Value
nn_type_short	1
random_number_type	5
random_seed mode 1	13242
random_seed mode 2	37198
number_of_elements	3
elements	H O Al
remove_atom_energies atom_energy H / E_h	-0.50498442
atom_energy O / E_h	-75.16628513
atom_energy Al / E_h	-243.03908176
energy_threshold	-0.215
bond_threshold	0.5
cutoff_type	1
use_short_nn	
global_hidden_layers_short	2
global_nodes_short	15 15
global_activation_short	t t l
test_fraction	0.1
check_input_forces	1.0
epochs	20
points_in_memory	20000
mix_all_points	
scale_symmetry_functions	
center_symmetry_functions	
fitting_unit	eV
precondition_weights	
use_short_forces	
optmode_short_energy	1
optmode_short_force	1
kalman_lambda_short	0.98
kalman_nue_short	0.9987
element_decoupled_kalman	
short_energy_fraction	1.0
short_force_fraction	0.025
weights_min	-0.5
weights_max	0.5
nguyen_widrow_weights_short	
short_energy_error_threshold	0.0
short_force_error_threshold	0.2
force_update_scaling	1.5
max_force	0.11
write_trainpoints	
write_trainforces	

Curriculum Vitae

Personal Data

Name	Martin Liebetrau
Date of Birth	March 30, 1994
Place of Birth	Witzenhausen, Germany
Nationality	German
Email	martin.liebetrau@gmail.com

Education

2013 – 2016	Bachelor of Science in chemistry at the Georg-August-Universität Göttingen Bachelor's thesis supervised by Prof. Dr. Alec Wodtke Title: Chemometrische Differenzierung mehrerer Mikroplastikpartikel mittels Raman-Spektroskopie (September 9, 2016)
10/2016 – 09/2018	Master of Science in chemistry at the Georg-August-Universität Göttingen Master's thesis supervised by Prof. Dr. Jörg Behler Title: Adsorption von Wasserstoff Atomen auf Aluminiumoxid (September 28, 2018)
11/2018 – 6/2023	Doctoral candidate in chemistry supervised by Prof. Dr. Jörg Behler in the Institut für Physikalische Chemie at the Georg-August-Universität Göttingen Title: Reactivity at Surfaces with High-Dimensional Neural Network Potentials (dissertation: April 30, 2023, oral examination: July 2, 2023)

Awards

10/2015 – 09/2016	Deutschlandstipendium
10/2017	Niedersachsenstipendium
SS 2018	Award for the best tutor in the Institut für Physikalische Chemie at the Georg-August-Universität Göttingen in the summer term 2018