

Improved *ab-initio* research of hydrogen coverage on RuO₂ and IrO₂



présenté par

Dipl.-Phys. Jan Kloppenburg

Promoteurs:

Professeur Geoffroy Hautier

Professeur Gian-Marco Rignanese

Jury:

Professeur Geoffroy Hautier

Professeur Gian-Marco Rignanese

Professeur Jean-Christophe Charlier

Professeur Sophie Hermans

Professeur Jin Suntivich

Professeur Philippe Vereecken

Professeur Bernard Nysten

Louvain-la-Neuve, Belgium, December 4, 2023

The usual approach of science of constructing a mathematical model cannot answer the questions of why there should be a universe for the model to describe.

Why does the universe go to all the bother of existing?

– *Stephen Hawking*

Für Birte, die Liebe meines Lebens

Für Robin, Du fehlst

Abstract

Despite promising developments of experimental techniques the molecular processes on the surface of catalysts are still not fully understood. Hydrogen atoms binding to the surface atoms are hard to detect with modern techniques because of the nature of hydrogen. In this work we will first gain insight into the hydrogen coverage and the atomic configuration on the surface of iridium oxide and ruthenium oxide by computing surface phase diagrams with *ab-initio* methods. With these findings we will move on towards explaining the improvement of catalytic reactions on the iridium surface when mixing in titanium to enhance the activity which is experimentally evident and will be elucidated by *ab-initio* methods. The computation of a cyclic voltammogram measurement by means of a cluster expansion in combination with machine learning will be presented in the last chapter which is also compared to experimental data. Finally a discussion about the results from this work, the context of their validity and also the shortcomings of the currently available methods is summarized with an outlook in the final paragraph.

Contents

1. Introduction	1
1.1. Energy production today	1
1.2. Water Electrolysis	2
1.3. State of the Art	9
1.4. Theoretical model and the real electrochemical cell . . .	11
1.5. Flow of this Work	12
2. Theoretical Background of Applied Methods	13
2.1. Density Functional Theroy	13
2.2. Grand Canonical Surface Phase Diagrams	20
2.3. From Statistical Mechanics to Cluster Expansion	24
2.4. Mathematical formulation of the Cluster Expansion . .	30
3. Surface states during CV measurements	45
3.1. Molecular Beam Expitaxy (MBE)	46
3.2. Cyclic Voltammetry (CV)	47
3.3. The influence of pH on the CV shape	52
3.4. Interpretation of the CVs	53
3.5. Surface Phase Diagram Definition	54
3.6. Chemical Potentials	54
3.7. Surface Structures	56
3.8. The IrO ₂ Surface	60
3.9. The RuO ₂ Surface	68
3.10. Adsorbed water on RuO ₂	72
3.11. Solvation energies	73
3.12. Results	76

4. Ti-IrO₂ alloy	79
4.1. Oxygen Evolution Reaction	79
4.2. Pure IrO ₂	80
4.3. Can we improve IrO ₂ by alloying Ti?	85
4.4. Experimental Data	86
4.5. Possible motifs for Ti-enhanced IrO ₂	86
4.6. Results	92
5. Hydrogen coverage of the IrO₂ Surface	93
5.1. Defining the Problem	94
5.2. Generating Structures for Training Data	96
5.3. DFT code details	96
5.4. Grand Canonical Phase Diagrams	102
5.5. Hydrogen chemical potential alignment	105
5.6. DFT Formation Energies	106
5.7. Clusters in the Expansion	111
5.8. Monte-Carlo Simulations	112
5.9. Adding Zero-Point Energies	115
5.10. Interpreting the Monte-Carlo Results	117
5.11. Results	125
6. Discussion and Perspectives	127
Appendix	CXLVII
I. DFT equations and background of methods	CXLVIII
II. CASM: Command line output	CLI
III. Monte-Carlo settings:	CLIV
IV. Basis set parameters and resulting fit for ZPE CE	CLVIII
V. Ti-alloyed host structures	CLIX
VI. Monte-Carlo Snapshots	CLXII

1. Introduction

1.1. Energy production today

Despite the known impact of human produced CO₂ on the atmosphere and the change in climate [1], in 2019 the bulk production of global energy still heavily relies on fossil fuels [2] as seen in left figure 1.1.1. Atmospheric CO₂ levels are on historic highs (right figure 1.1.1, from Ref. [3]) and keep rising and despite the Paris climate accord and are predicted to increase even further [4, 5].

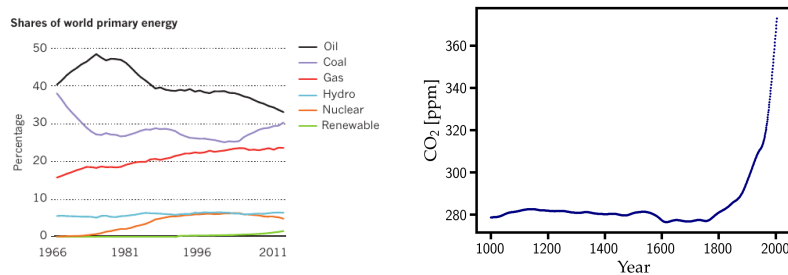


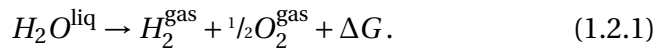
Figure 1.1.1.: Left: Energy production and usage from different sources, adapted from [2]. **Right:** CO₂ levels from ice core data from the antarctic, published in [3].

To overcome the challenge of a continuously growing energy demand and climate changes that can potentially make large areas of the planet uninhabitable we need to change the way energy is produced. Many alternative production methods exist and are already employed to-

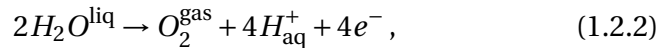
day. The most prominent today are solar, wind and hydro based production facilities. The main challenge with the renewable energy technologies lies in the storage of energy. In contrast to fossil fuel based energy production, renewable sources are not always available when energy demand is high but might have production peaks during times when demand is low. For this reason the need for large scale storage facilities is greater than ever. There is a variety of storage technologies available today ranging from mechanical [6], chemical [7], thermal [8] to electrochemical [9] storage technologies. Each of these faces its own challenges in efficiency and technological development and not every large scale solution is applicable for a given location. Hydroelectric storage facilities need geological formations suitable to house the required basins and pumping installations. If these conditions can not be met a different storage technology or location has to be chosen. In addition to the requirements of the storage solution, different technologies offer specific ranges of applicability shown in figure 1.1.2. The final choice of storage technology will be determined by a beneficial combination of many factors.

1.2. Water Electrolysis

In 1789 Adriaan Paets van Troostwijk and Johan Rudolph Deiman first realized the electrolysis of water [11]. The complete reaction is written down to



The energy ΔG required for this reaction to proceed is calculated to $\Delta G = 237.2 \text{ kJ/mol} = 2.46 \text{ eV/mol}$ [12]. This energy is for the complete reaction involving two electrons thus the energy needed per electron is 1.23 eV. The complete water splitting reaction is commonly separated into two half-cell reactions, written for the anode oxidation



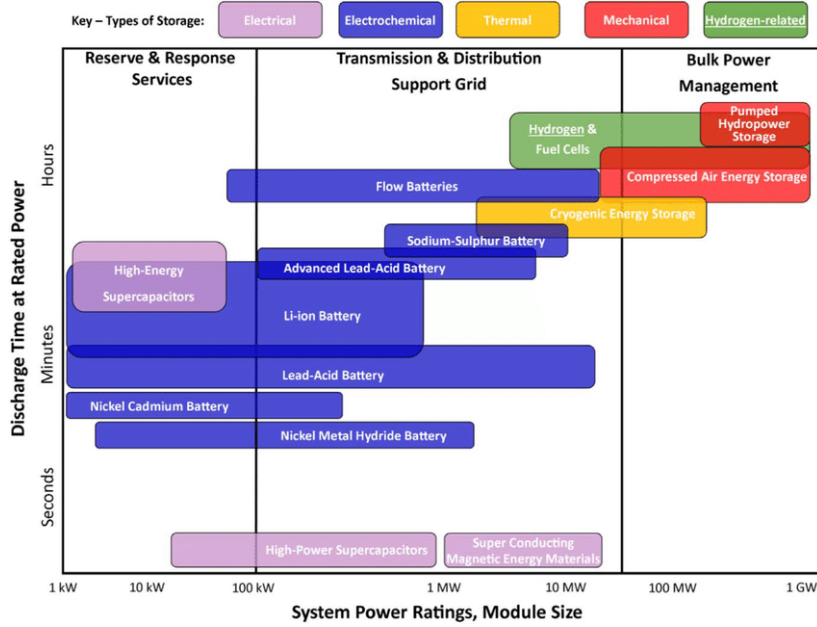
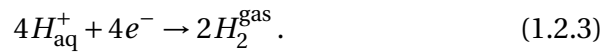


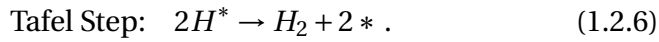
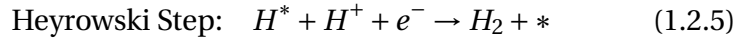
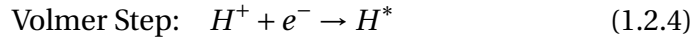
Figure 1.1.2.: Capacity vs Discharge comparison of current available technologies, from Ref. [10].

and the cathode reaction of hydrogen



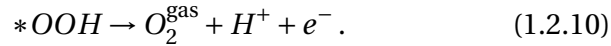
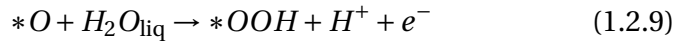
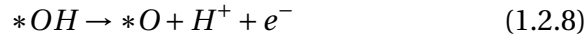
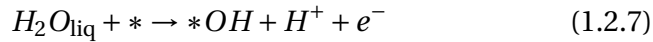
From equation (1.2.2) we know that the complete splitting involves four electrons and the total energy needed to produce 2 H₂ and O₂ from 2 H₂O is 4.92 eV. For the theoretical description of the hydrogen evolution reaction (HER) and the oxygen evolution reaction (OER) we need a reaction cycle that expresses the single reaction steps in equations. While the reaction equation for water splitting is rather simple the reaction itself involves a four-electron process until the products are formed. It is much more complicated than the HER which is de-

scribed by



The $*$ in these equations represents an open catalytic surface site. The overall process always includes the Volmer step which describes the adsorption of a proton on the catalyst surface from which it takes an electron. The second step then is the desorption described by either equation (1.2.5) or (1.2.6). The Heyrowski step describes recombination of a surface adsorbed hydrogen with a proton from solution and an electron from the catalyst. The Tafel process describes the recombination of two surface adsorbed hydrogen atoms which combine to form H_2 and then desorb.

To describe the OER we need four equations:



Equation (1.2.7) refers to a water molecule from which a proton and an electron are split off. The electron travels through the anode while the proton travels through the electrolyte to the cathode to either adsorb in a Volmer step (equation (1.2.4)) or recombine with an existing adsorbed hydrogen via Heyrowski (1.2.5) or Tafel (1.2.6) processes. In all four equations (1.2.7)-(1.2.10) the electron-proton pair behaves in this manner. The remaining OH adsorbs at a suitable surface site $*$. In the second step (equation (1.2.8)) the H is stripped off the adsorbed $*OH$ leaving behind only the adsorbed oxygen $*O$. In equation (1.2.9) another water molecule from the electrolyte is split and the $*OH$ adsorbs onto the already present $*O$ on the surface, forming $*OOH$. For many catalyst materials, the third step, the $*OOH$

formation, is the rate limiting step in the OER. In equation (1.2.10) the molecular oxygen desorbs from the surface site after the splitting of an electron-proton pair, leaving behind the empty surface site so the cycle can start again. Neither the HER nor the OER reactions proceed without external potential. The theoretical value for each single reaction step in equations 1.2.7-1.2.10 is 1.23 V and a perfect catalyst would facilitate each step at this energy. Any potential that has to be applied above the 1.23 V is called overpotential and can be regarded as energy loss that can not be recovered. While the HER proceeds with low overpotentials on different catalytic materials [13, 14], the oxygen evolution has considerable overpotential on the best materials today [15, 16]. Further reduction of OER overpotential is active research and requires the design of new materials. An approach to the design of specific OER materials is presented in this work in chapter 4. Today there are many different specially designed catalysts for specific chemical reactions [17]. The industrial gold standard for water splitting is the performance of IrO_2 and has been investigated extensively [16]. Iridium metal is binding the oxygen just enough to make the reaction possible but also not too strong to let the reaction product O_2 desorb from the surface. The efficiency of catalysts is not only defined by the metal used in the production process but also by other elements present, i.e. doped surfaces, impurities, defects or structural features of the surface like islands or steps and the surface orientation itself [17]. Designing a new and highly efficient catalyst is thus a complex problem which can not be solved by addressing only a single issue, i.e. the metal. To create new catalysts, we need to gain a deeper understanding into the underlying chemical processes that take place before and during the reaction.

1.2.1. Relation of adsorption energies to reaction barriers and the volcano plot

Catalytic reaction on surfaces are dominated by the kinetics of the reaction. While we are not computing kinetic barriers from DFT to calculate the theoretical overpotentials for our metal surfaces we assume the Brønsted – Evans – Polanyi (BEP) [18, 19] scaling relation holds. It states that the kinetic barrier for a reaction scales linearly with the adsorption energies. Kinetic barriers can be computed from DFT but they need much more computational time than adsorption energies only. Assuming the BEP relation holds for our case we can therefore derive overpotentials from adsorption energies. When comparing the results from these computations to experimental values we see that the trends are covered correctly. Materials with low theoretical overpotentials show experimentally higher activities than materials with higher theoretical overpotentials. To visualize this, a linear relationship enables us to draw lines on which all transition metals catalysts should fall, theoretically. These lines are stemming from two dominating parts during the OER. On the left side there is strong oxygen binding resulting in $*O \rightarrow *OOH$ being the rate limiting step while on the right side we have weak oxygen binding resulting in $*OH \rightarrow *O$ being the rate limiting step. With computed values for the theoretical overpotential η_{th} and the descriptor $\Delta G_{*O} - \Delta G_{*OH}$ for the x-axis we can add a data point for each transition metal surface we compute. This is visualized with in a volcano plot shown in figure 1.2.1. The most active compounds are located near the top of the volcano. From experiment, we know that IrO_2 and RuO_2 are highly active OER catalysts and both are found near the top. We also know that TiO_2 is regarded as inactive for OER and consequently we find its value on the right lower side indicating too weak binding of oxygen and an overpotential of about 1.2 V. In general this plot gives a good view of the materials and their respective overpotentials.

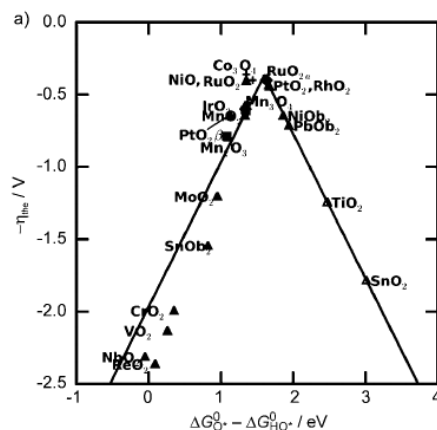


Figure 1.2.1.: Volcano plot for different metal oxides from Man [20]. Compounds with highest activities are near the top of the volcano.

1.2.2. Catalysts

The water splitting reaction 1.2.1 is generally facilitated by a catalyst. The efficiency of the reaction largely depends on the choice of catalyst material while the most active materials today are IrO_2 and RuO_2 [16, 21]. Many efforts are made to reduce energy losses during water electrolysis and improve the catalyst by modification of the structure, different surface orientations and compositing of highly active base materials with selected metals [13, 22, 23]. This effort is a slow and in many cases very expensive process when done experimentally. For a reliable classification of the material one usually needs well defined surfaces which are hard to produce. A good way to produce these surfaces is by molecular beam epitaxy (MBE) as one ends up with a well defined surface orientation and composition. The catalysts investigated experimentally in this work were grown by MBE and the theoretical model was adapted from the orientation of the experimental

sample. This ensures that the correct surface model is being investigated both theoretically and experimentally and that the results are directly comparable.

1.2.3. The active surface

Before we can begin the theoretical description of catalytic materials during electrochemical reactions we need to understand the surface of the catalyst. For this purpose the choice fell towards the widely studied transition metal oxide IrO_2 [16, 21]. Here we specifically use a surface orientation compatible with experimental measurements and we will compare our computational data to the experimental data. The catalyst was experimentally grown by molecular beam epitaxy [15, 24] and presents an ideal and pristine surface for both experimental measurements and theoretical description. The orientation of the crystal is the (110) direction of the rutile structure. In the theoretical model we will not consider any defects on the surface, i.e. no missing metal or oxygen atoms are assumed. At this time we do not have an experimental classification of the surface defects for the models we investigate. The surface configuration depending on externally applied potentials will be investigated to understand the properties of the catalyst under working conditions at the atomic level. In the first part both RuO_2 and IrO_2 will be investigated using surface phase diagrams. Later we will only look at IrO_2 for two reasons, first it is experimentally known that RuO_2 dissolves by a factor of magnitude stronger at 1.2 V than IrO_2 in acidic conditions [25] and second, the computational cost of the applied method did not allow to study both, thus the more stable material was chosen.

1.2.4. Splitting water

During the electrolysis of water gaseous hydrogen and oxygen are produced. Both gases can be stored in tanks, the oxygen can be re-

leased into the atmosphere if needed. The stored hydrogen can later be reduced to water in a fuel cell releasing the stored chemical energy and producing water with atmospheric or stored oxygen. We have to keep in mind that for the production of molecular hydrogen for electrochemical energy storage both half cell reaction have to proceed evenly. Only the hydrogen recombination proceeds with low overpotential today, the oxygen production is sluggish and thus hinders the whole reaction to proceed easily. Although the focus of this work is the cathodic production of hydrogen for storage and release of energy the potential applications of electrochemistry will not be neglected here. The central functional element in electrochemistry is the catalyst. Special designs of catalytic materials allow for high selectivity in the facilitated reactions and thereby allow a wide array of potential applications. An exemplary cycle of possible applications is shown in figure 1.2.2. The area of catalyst design and selectivity is active research today. Mobility of adsorbed species and the number of active catalytic sites are main factors of influence to the catalytic activity. The choice of materials will in general influence the selectivity of the active catalyst such that every targeted reaction has its own facilitating catalyst. In addition to these challenges the development of non-toxic materials from earth-abundant metals is of central significance [22, 23]. As the main focus of this work we regard the innermost cycle describing the hydrogen production and reduction from figure 1.2.2 of central importance.

1.3. State of the Art

The current knowledge of the atomic surface structures of hydrogen is limited. This is mainly because experimental measurements of hydrogen are difficult, especially *in-operando*. Therefore, only indirect measurements of hydrogen activity are provided with the most common technique being the cyclic voltammetry. It will be explained in

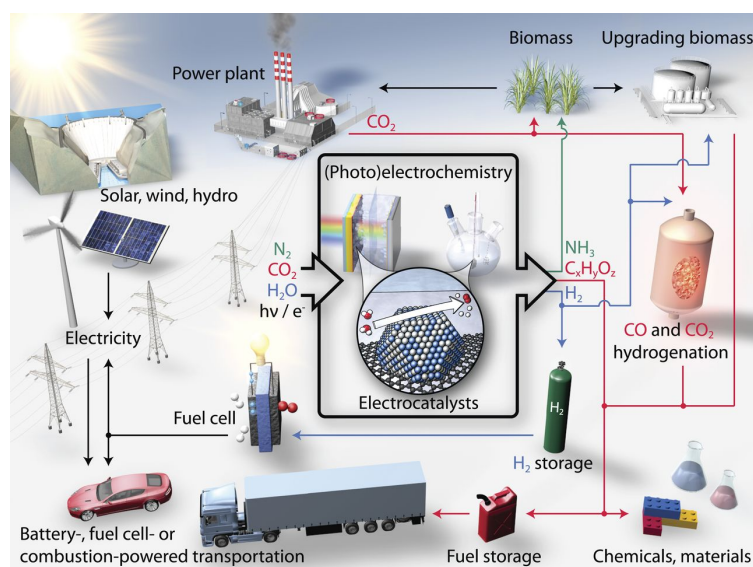


Figure 1.2.2.: Overview of processes and applications around a central electrochemical catalyst, from Ref [13]. This work is related to the innermost circle of electrochemical hydrogen production but the potential applications of electrochemistry are abundant.

detail in chapter 3 when results from these measurements are presented. Another limiting factor in the description of hydrogen coverage on the surface is the computational cost in the description of large surfaces. Common approaches therefore concentrate on small surfaces with only limited possibilities for hydrogen configurations and do not provide enough information for the reconstruction of a computational cyclic voltammetry for comparison to experimental data. With improved understanding of the surface processes involved before and during the oxygen evolution reaction we will be able to discover the limiting factors of the reaction which enables the design of more active catalyst surfaces.

1.4. Theoretical model and the real electrochemical cell

Electrochemical cells are complex reactors with many factors at play. Aside from the metal electrodes that are not the perfect structures we model in theory there is also the electrolyte. For the storage of energy in hydrogen the main portion of the electrolyte is water. Other parts are usually additives like KOH, LiOH or NaOH to adjust pH towards more basic levels or H_2SO_4 to adjust towards acidic conditions. Large ions in the electrolyte stemming from the pH level adjustment are present in the real system but are not modeled in our theory. The computational cost to include realistic concentrations is too high. Furthermore, the real water structure in contact with a charged surface forming polarized layers is very complicated and changes with the applied potential, too [26, 27]. This can give rise to a pseudo-capacitance within the Stern and Helmholtz layer [26]. These effects are not included in our model for the reason of computational cost. Additionally, we assume the protons are travelling instantly through the electrolyte for recombination at the cathode which in the real system is defined by diffusion. The process is not instantaneous but takes a finite time in the experimental cell. Our theoretical model does not include any diffusive processes. The model focuses on the theoretically accessible parts of the electrochemical cell system. We model the atomic structure of hydrogen on the surface of the anode where the oxygen evolution reaction takes place. We do not include any implicit or explicit water structure above the surface and we do neglect the contributions of anions or cations in the electrolyte from additives to adjust pH levels.

1.5. Flow of this Work

The main focus of this work is to model the surface before the oxygen evolution reaction at the atomic level. This will grant new understanding of the catalytic processes at the atomic level. Computational tools like density functional theory will serve as the main utility to compute the properties of relevant surface structures. From these results the derivations towards surface states will be made while an external potential is simulated. The results from these models are compared to experimental results directly. We will move from relatively small surface structure models towards the description of a large surface from which we can then derive quantities from statistical mechanics theory. The modeling of the surface within statistical mechanics methods is possible by employing the cluster expansion method. This method yields further insights into the electronic structure effects that determine the shape of the experimental measurements. We will compare these results directly to the experimental results as we have done before for the smaller surface models. Then we move on from processes before the OER to the reaction itself as we have now understood the properties of the surface up to OER potential. The surface we have investigated up to this point will then be our starting point in chapter 4. The computational description of the OER is presented but also improvement in activity has been shown experimentally by doping the surface with titanium. This seems surprising at first because TiO_2 in the same rutile configuration as IrO_2 is regarded as inactive with respect to the OER. This can immediately be seen in the volcano plot in figure 1.2.1. With the acquired knowledge from this work we are able to explain the increase in activity for this new material.

2. Theoretical Background of Applied Methods

2.1. Density Functional Theory

Density Functional Theory (DFT) is a popular and very successful tool in theoretical chemistry and physics today. The roots of DFT lie in the Hohenberg Kohn theory [28] which was one year later developed further by Kohn and Sham to the DFT [29]. The theory of DFT is described in many modern textbooks such as Refs [30, 31, 32]. Today there are many different DFT codes employing different mathematical methods. The perhaps most widely used DFT code is VASP [33, 34]. For a detailed account of the mathematics and technical implementation the interested reader is referred to the cited resources. In short, the Kohn-Sham equations are solved in a discrete basis set and the resulting eigenvalues and eigenvectors are regarded as single-particle energies and their respective wave functions. The Kohn-Sham equations are

$$v^{\text{eff}}(\mathbf{r}) = v(\mathbf{r}) + \frac{e^2}{4\pi\epsilon_0} \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r}' + \frac{\delta E^{\text{xc}}[n]}{\delta n(\mathbf{r})}, \quad (2.1.1)$$

and

$$\left(-\frac{\hbar^2}{2m_e} \nabla^2 + v^{\text{eff}}(\mathbf{r}) \right) \psi_{o_i}(n(\mathbf{r})) = \epsilon_{o_i} \psi_{o_i}(n(\mathbf{r})), \quad (2.1.2)$$

and we can express the density as

$$n(\mathbf{r}) = \sum_{i=1}^N |\psi_{o_i}(\mathbf{r})|^2. \quad (2.1.3)$$

From the solution of these equations we obtain the single-particle energies and the wave functions for each electron and with these we can compute further properties we are interested in.

2.1.1. Structure Relaxation

From the solution of the Kohn-Sham equations one obtains not only the density $n(\mathbf{r})$ but also the wave functions used during the process. From these we can compute the forces acting on each ion in the structure. When creating structures from scratch the user has an initial idea of the geometry configuration to start with. This structure then needs to be optimized with respect to the forces acting on the ions. The algorithm employed in many DFT codes is the BFGS method [35] or improved variants of it [36]. The BFGS method belongs to the quasi-Newton methods. The computed forces acting on the ions are applied to their respective center of mass, the ions are moved accordingly and the SCF cycle restarts. This loop continues until user-defined criteria are met. A common target value used to define a converged structure is the remaining maximum force component. If this force is below a certain value, we commonly use 10^{-2} eV/Å in this work, the structure is regarded as converged.

2.1.2. Exchange Correlation Functionals

In the work of Kohn and Sham [29] the expression for the energy E depending on the density of the electrons n is given by

$$E[n] = T_s[n] + \int v(\mathbf{r}) n(\mathbf{r}) d^3\mathbf{r} + E_{\text{Ha}}[n] + E_{\text{xc}}[n]. \quad (2.1.4)$$

All but one term in this equation are known and can be computed. The remaining unknown is the exchange correlation functional $E_{\text{xc}}[n]$. For this functional there are different approximations available which will be briefly described below. Common shortfalls of exchange correlation parametrizations are

- van-der-Waals and hydrogen bonding
- highly correlated materials
- excited states

The development of these functionals is ongoing active work.

Local density approximation (LDA)

The LDA was first developed and originally only thought to be useful for metals and periodic bulk structures but as it turned out this rather drastic approximation is surprisingly good. The exchange-correlation energy is defined by

$$E_{xc}^{\text{LDA}}[n] = \int n(\mathbf{r})\epsilon_{xc}(n)d\mathbf{r}. \quad (2.1.5)$$

The approximation is now done by inserting a specific $\epsilon_{xc}(n)$ that is a solution of the homogeneous electron gas. One can also include the spin as further input and expand the functional into local spin density approximation (LSDA)

$$E_{xc}^{\text{LSDA}}[n_{\uparrow}, n_{\downarrow}] = \int n(\mathbf{r})\epsilon_{xc}(n_{\uparrow}, n_{\downarrow})d\mathbf{r}. \quad (2.1.6)$$

In the LDA case, the exchange part is known analytically from the work of Bloch [37] and Dirac [38], and popular parameterizations of LDA correlation functional are from

- Wigner [39]
- Perdew and Zunger [40]
- Perdew and Wang [41]
- von Barth and Hedin [42]

Generalized gradient approximation (GGA)

An improvement over the LDA would of course be to not only include the spatial coordinates but also the charge variation in the description, thus making the density a function of both which then is to be called a generalized gradient approximation. The work by Langreth and Mehl [43] shows how to construct working GGA functionals.

$$E_{xc}^{\text{GGA}}[n] = \int n(\mathbf{r}) \epsilon_{xc}(n, \nabla n) d\mathbf{r}. \quad (2.1.7)$$

As in the LDA case the approximation is now done by inserting a parameterized $\epsilon_{xc}(n, \nabla n)$ and one can as well include the spin as further input and expand the functional into

$$E_{xc}^{\text{GGA}}[n_{\uparrow}, n_{\downarrow}] = \int n(\mathbf{r}) \epsilon_{xc}(n_{\uparrow}, n_{\downarrow}, \nabla n_{\uparrow}, \nabla n_{\downarrow}) d\mathbf{r}. \quad (2.1.8)$$

Some popular parameterizations of GGA functionals are from

- Perdew, Burke and Ernzerhof (PBE) [44, 45]
- Hammer, Hansen and Nørskov, revised PBE (RPBE) [46]
- Becke, Lee, Yang and Parr (BLYP) [47, 48]
- Schultz, Zhao and Truhlar [49]

Hybrid Functionals

Hybrid functionals are approximations that combine a percentage of exact exchange from Hartree-Fock theory with exchange correlation from other empirical or *ab-initio* sources [50, 51, 52, 53]. The exact exchange is mixed by linear combination and fitted to either theoretical or experimental data. They are computationally more expensive and describe certain properties as band structure more accurately. In this work we will not make use of this class of functionals.

2.1.3. Our choice of functional - RPBE

We use the RPBE functional [46]. It describes the chemisorption processes for molecular adsorption of O, CO and NO on metal surfaces more accurately than the regular PBE, in particular the oxygen binding to surfaces is described more accurately (as explained in Ref. [46]) which is of central importance in this work.

2.1.4. Van-der-Waals Interactions

Van-der-Waals interactions (vdW) are not included in the DFT theory. Long or short range electrostatic phenomena are in general not included in DFT. These phenomena are particularly relevant when partially charged ions are considered or when strong electronegative ions like oxygen bind with hydrogen to form water. Considering liquid water molecules it has been shown that van-der-Waals corrections do improve the mobility of water simulated by *ab-initio* when comparing to neutron diffraction experiments [54]. Also the combination of hydrogen networks and van-der-Waals effects are responsible for the maximum density of water and the negative volume of melting [55]. To include van-der-Waals effects in DFT there are two principle ways:

- Use an exchange correlation functional that includes these effects [56, 57]. This method is in principle exact concerning the van-der-Waals long range effects but due to its nature very expensive. Only small systems can be computed in realistic time frames and thus for many problems that can be solved with DFT and GGA the use of a van-der-Waals-functional greatly increases computational cost.
- After the SCF cycle (figure I.1 appendix) is converged one applies a van-der-Waals correction based on the computed electronic density. This method is not exact in terms of van-der-Waals corrections but it is cheap to apply and systems that can

be computed using regular GGAs can also be computed with GGA+vdW. One specific method on this base was developed by Tkatchenko and Scheffler [58] and when applied will be marked RPBE+vdW^{TS} from now on throughout this work.

In this work we are not investigating the structure of water for which van-der-Waals interactions are important for an explicit molecular description. Instead we are computing hydrogen coverage on an oxygen covered metal-oxide surface. Polarization from hydrogen bonding and van-der-Waals effects are not the dominant contributions but the effects are not neglected by including the van-der-Waals correction introduced by Tkatchenko and Scheffler [58]. In chapters 3 and 5 we employ the RPBE+vdW^{TS} method. In chapter 4 we do not include vdW^{TS}. This is because the literature we use as comparison does not include this correction. For these computations we use the plain RPBE functional.

2.1.5. Implicit Solvation Model

When running DFT computations on surface structures one usually has a region that is vacuum as shown in figure 2.1.1 left. This situation is in most cases not a description of the actual experiment. In this work we are considering a surface in contact with water. An explicit model of all water molecules is in general too expensive. An approximation of the liquid water that is not directly involved in the atomic processes at the surface is a reasonable measure. The implicit water model presents itself as an electrostatic background that has the dielectric constant of water $\epsilon_r=78.4$ at $T=298K$. With this approach one reduces the total number of atoms in the DFT simulation drastically and still captures some effects of a water layer above the simulation surface. In electrostatic models one does expect stronger effects with higher polarizability of the material. This is represented within the implicit solvent models. There are many different implicit models available in different DFT codes. Starting from a simple electrostatic

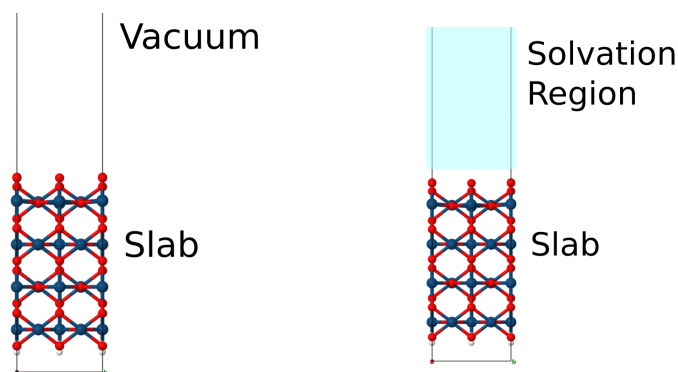


Figure 2.1.1.: Example of a slab in vacuum and with implicit solvation model, **Left:** the slab without model with adjacent vacuum, **Right:** the slab modeled with implicit solvation region.

background as just described more complex models have been developed to account for more complex physical models. A short list of available models:

- VASPsol [59, 60]
- jDFTx implementations:
 - Linear PCM [61, 62, 63]
 - Nonlinear PCM [61, 64]
 - SaLSA [64]
 - ClassicalDFT [65, 66]

The experience during this work has shown that the solvation effects are of similar magnitude as the van-der-Waals effects. The computational cost however is greater than the inclusion of the van-der-Waals effects.

2.2. Grand Canonical Surface Phase Diagrams

2.2.1. Surface Phase Diagram Definition

One way to classify different surface structures of similar materials are surface phase diagrams. They are successfully employed in many different areas of materials science and surface science. The mathematical definition of the surface energy reads

$$\Gamma = \frac{1}{A} \left(E_{\text{slab}}^{\text{DFT}} - \sum_i n_i \mu_i \right), \quad (2.2.1)$$

wherein A denotes the surface area of the slab, $E_{\text{DFT}}^{\text{slab}}$ is the computed total energy of the slab while n_i and μ_i represent the number of constituents and their chemical potential, respectively. The area of the slab is included to be able to compare different surface sizes to each other.

2.2.2. The reference Electrode: RHE vs SHE

The terms RHE and SHE are commonly used in cyclic voltammetry as refer to the reference electrode. RHE notes the Reversible Hydrogen Electrode while SHE stands for Standard Hydrogen Electrode. The scale is in both cases linear compatible but carries an offset when converted. The conversion reads

$$U_{\text{SHE}} = U_{\text{RHE}} + k_B T \ln(10) \cdot \text{pH}, \quad (2.2.2)$$

where we note k_B as Boltzmann constant and T as temperature. The pH level is a constant shift in this conversion only, it enters the equation as a linear additive term that scales with the pH value. The experimental data we use to compare our results to is measured with the RHE and thus in this work the reference electrode is chosen to be RHE.

2.2.3. Construction

The construction of a surface phase diagram is straightforward. Once the chemical potentials are defined we can compute Γ for each surface structure at two external potential points and draw a line connecting these. Furthermore we can choose a reference state by defining a number of constituents n_i . Then the values in the plot become

$$\Gamma_{\text{Plot}} = \Gamma_{\text{Surface}} - \Gamma_{\text{Reference}} \quad (2.2.3)$$

and the plots are easier to interpret. In general, the reference is chosen by the user to represent a defined starting point of a series of computations or as a boundary of chemical potentials. In our case we choose a distinct surface configuration as discussed in chapter 3.

2.2.4. Surface Structures

We have seen a slab in side view in figure 2.1.1. The surface we are investigating in this work is the (110) face of the rutile bulk shown in figure 2.2.1. From the bulk rutile structure we generate the (110) surface by cutting along the (110) plane of the bulk. Afterwards a rotation is performed to have the vacuum vector pointing in z-direction which is a technical requirement for some DFT codes and does not change computational results. We took a four tri-layer slab for our simulations. From literature [67] we know that this structure can be regarded as producing converged results in terms of adsorption energy computation. From the same literature we also adopt the k-point sampling of $8 \times 8 \times 1$ for the 1×1 surface cells.

2.2.5. Dipole Correction

All surface structures computed in this work are periodic structures. Since we are investigating surfaces the unit cell contains the structure and a vacuum region. To avoid interactions of the lower surface

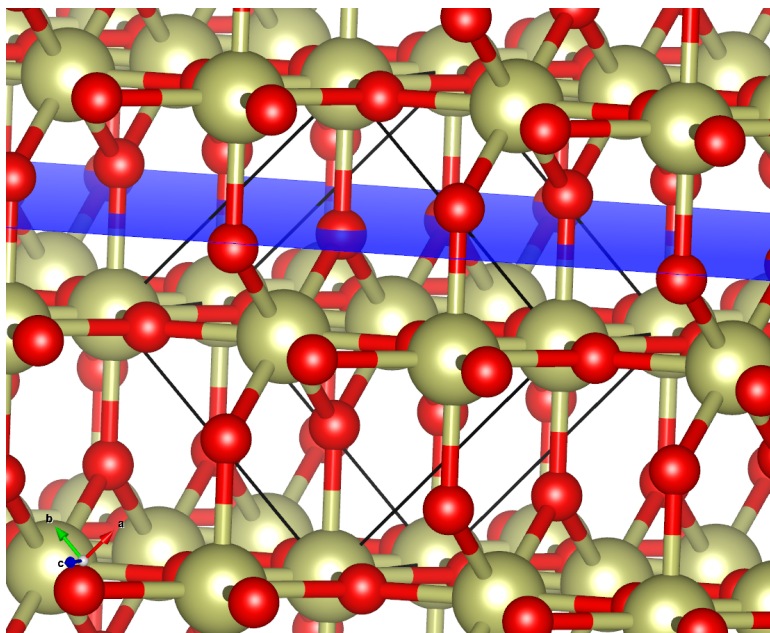


Figure 2.2.1.: IrO_2 in rutile structure. Unit cell is indicated in black lines and the 110 plane in blue. Slicing the model at the blue line yields the (110) surface we are modeling.

with the upper surface side the dipole correction [68] was developed. By computing an electrostatic potential step in the vacuum region a compensation of possible surface dipoles is achieved. This way, unphysical interactions are prevented. Since we employ two different DFT codes that use different implementations we discuss both here. In FHI-aims [69, 70, 71] a dipole correction is essentially unnecessary. The code uses a localized orbitals implementation and because of it the computational cost does not increase when adding vacuum to the cell. In this work all cells computed with FHI-aims have a vacuum vector of $100\text{\AA}$. This ensures a minimum vacuum distance of at

least $80\text{\AA}$ and surface-surface interactions are not present. In VASP [33, 34] this scenario is different concerning the computational cost. VASP is a plane-wave code employing pseudo-potentials [72, 73]. In this implementation the computational cost is negatively impacted when vacuum is added to the simulation cell. To keep the cell as small as possible without interaction errors from polarized surfaces the dipole correction is important. For this reason the dipole correction is regularly employed when surface simulations are computed in plane-wave codes. We use it as implemented in VASP [68]. A test for the minimal vacuum length was performed and the result is shown in figure 2.2.2. From figure 2.2.2 we can see that the dipole correction

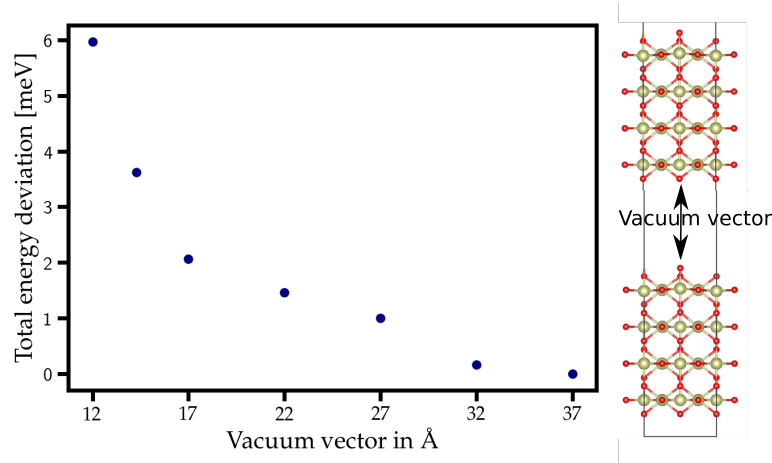


Figure 2.2.2.: Vacuum distance convergence test. At a separation distance of $12\text{\AA}$ the error due to slab-slab interaction is below 6 meV . A separation of $37\text{\AA}$ is regarded as reference for a non-interacting systems.

decouples the periodic images very well. At the smallest distance of $12\text{\AA}$ there is only a 6 meV deviation from the as converged regarded result at $37\text{\AA}$. When we use VASP during this work the dipole correc-

tion will be employed by default. A minimum length of 15Å of vacuum will be ensured for all computations.

2.3. From Statistical Mechanics to Cluster Expansion

In grand canonical statistical mechanics the central quantity is the partition function defined as

$$\Xi(\mu, V, T) = \sum_i^{\infty} \exp\left(\frac{\mu N_i - E_i}{k_B T}\right), \quad (2.3.1)$$

where the chemical potential is noted as μ , the number of particles in the microstate i labeled N_i and their energy noted as E_i . Noted are also volume V , temperature T and the Boltzmann constant k_B . From this general quantity we can derive the grand canonical potential as

$$\Phi_G = -k_B T \ln \Xi = \langle E \rangle - TS - \mu \langle N \rangle, \quad (2.3.2)$$

where S denotes the entropy. The expressions $\langle E \rangle$ and $\langle N \rangle$ denote the quantum mechanical expectation values of E and N , respectively. In general the sum over i in equation (2.3.1) is an infinite sum so that the equation becomes impractical to solve for any computer implementation. In principle one could sum up a random number of terms chosen by some criteria but there is no way to tell whether this procedure has a convergence behavior or if this is reached. So without a concrete way to a decision about the completeness of the expression there is no criteria to decide if enough terms are included in the partition function to properly describe the system in question. Computationally this problem becomes unsolvable very quickly and has thus no practical application in theoretical quantum chemistry. One more problem in this description is the missing convergence criteria. If we had some way to tell whether we have enough terms included

to make an error approximation about the remaining terms we would be a step ahead. There is a way to achieve this by reordering the total energy expression. In the current form in equation 2.3.1 the ordering is done by enumerating the microstates. If one can reorder the expressions in terms of particle-particle interactions it would be beneficial for two reasons:

- (a) Interactions of short range are in general stronger than long range interactions
- (b) Interactions of $n + 1$ particles are usually weaker than interactions of n particles

Both points are steering towards an applicable convergence criteria. Statement (a) suggests that there might be a range that one can choose up to which specific interactions are included. Above that range the interactions are neglected. Statement (b) suggests that there is a maximum number of particle-interactions that need to be included (n) and interactions of $(n + 1)$ particles can be omitted. With these two statements there is now a mathematical way to achieve convergence computationally. Mathematically the formulation of a cluster expanded property $f(\vec{\sigma})$ reads

$$f(\vec{\sigma}) = \sum_{\alpha} j_{\alpha}(\mu) \phi_{\alpha}^{\mu}(\vec{\sigma}_{\alpha}), \quad (2.3.3)$$

wherein the $j_{\alpha}(\mu)$ represent the effective cluster interactions (ECIs), $\phi_{\alpha}^{\mu}(\vec{\sigma}_{\alpha})$ denote the cluster functions and α enumerates the clusters included in the expansion. In principle one can expand material properties (i.e. density of states or band structure) in this formalism but we will focus on the total energy of the compound. This will be formulated mathematically in more detail in the next section. Each cluster has an ECI associated to it and while there are many combinations of cluster geometries there are also many ECIs that are equal because

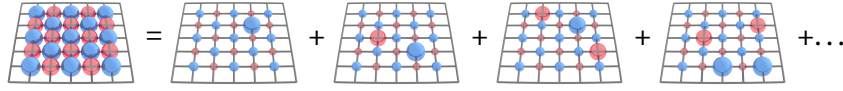


Figure 2.3.1.: Schematic display of the cluster expansion. The complete system on the left is represented as the sum of interactions of particle configurations. The first term are single particles, the second one pair interactions, the third term represents triplet interactions and the fourth quadruplets. In this schematic example the infinite sum is truncated at quadruplets. In the actual cluster expansion there are more than one single type of each cluster family represented, i.e. there is more than one pair contributing to the system representation, as is the case for all higher terms than singlets.

of the geometry of the cluster. The only remaining terms for the total energy evaluation then are the symmetrically distinct clusters and their ECIs.

In general, the interactions between particles are the quantity that is of central interest in the cluster expansion. An example is a binary alloy. Suppose we have a 2D square lattice of which each point can be occupied by either A or B. For this example it is not relevant whether A and B are atoms or one of them is a vacancy. Depending on how the lattice is populated we end up in different possible configurations, i.e. AA, AB, BA, BB for pairs. For triplet combinations we get AAA, AAB, ABA, BAA, ABB... Each single cluster has its own effective cluster interaction (ECI). These are labeled $j_\alpha(\mu)$ in the general equation 2.3.3. In principle there is no limit to the number of interacting particles included in the clusters of this expansion and the relating term is a basis function written as $\phi_\alpha^\mu(\vec{\sigma}_\alpha)$ in equation (2.3.3) as part of an infinite complete basis set. These cluster functions will not be given due to their complex mathematical nature. Much like when

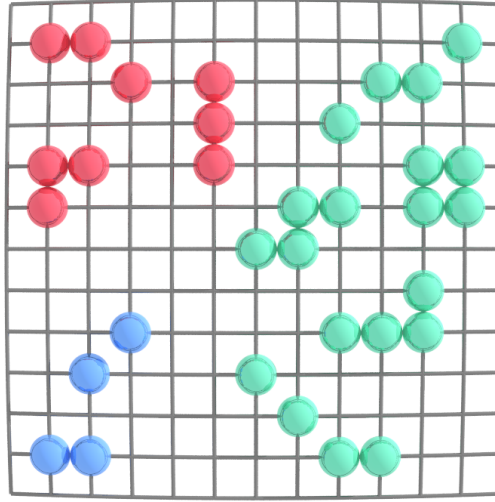


Figure 2.3.2.: Some possible pairs (blue), triplets (red) and quadruplets (green) occurring in a general cluster expansion. This is a general representation of possible configurations only. The actual clusters included in an expansion depends on user-defined parameters. In general the interacting particles do not have to be direct neighbors but can have some distance between them.

presenting the results from DFT computations one does not present the approximated basis set (aka wave functions) when using the total energy of a computation. The criteria used to construct the basis set will be given, i.e. a plane wave cutoff, while for the cluster expansion the number of included clusters will be given among the set of parameters used to represent the expansion. In practice one terminates this infinite series of basis functions at four or five particles and with that approximates the infinite series by a finite sum. From these

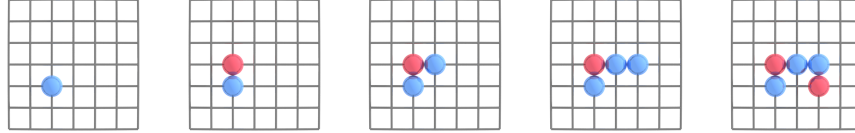


Figure 2.3.3.: Theoretically growing a binary alloy from single atoms as example. During this process the interactions between the particles are computed and used with the corresponding geometry as input for a cluster expansion. During this process the energy hierarchy of possible configurations is mapped. Note that the addition of particles is not necessarily ordered but can be in a rather randomized fashion because information about favorable configurations is not available.

ECIs we can determine the most favorable configuration occurring in the material, i.e. if ABA is more favorable than others this configuration will dominate the lattice. Depending on the differences of the ECI to other configurations there will be a portion of the lattice that is populated by different configurations. Depending on the symmetry of the underlying lattice the clusters and their respective ECIs will be reduced in numbers. If the nearest neighbor interaction of AB and BA are equal so are their ECIs. The cluster functions will include the symmetry information for this configuration. So what we have achieved by learning about the interactions of clusters is that we can now describe the lattice population in terms of cluster configurations. This enables the computation of total energies for large structures that can not be reasonably computed with DFT today. However, the termination criteria for the infinite cluster function sum is not given as a quantity that applies to every system equally. The number of particle interactions from which the clusters are constructed (pairs, triplets, ...) needs to be determined for each system individually. As

a general idea, think of an ionic solid like NaCl. In this case there will be strong pair interactions in the ECIs while triplets may already be weak. Thinking of a strongly covalent system like SiC the interactions of triplets and quadruplets will be a lot more important. So the basis set of cluster functions and the number of clusters needed to get a suitable description for each of these materials will be very different because of the physical properties of the materials. The practical fitting procedure is an initial guess from physical properties and trial and error. From a well fitted cluster expansion we can compute total energies of a structure by making use of a model hamiltonian constructed from the ECIs and their cluster functions.

2.4. Mathematical formulation of the Cluster Expansion

The cluster expansion theory [74] implemented in CASM[75] stems from ideas published in 1970. Further developments were published in Ref [76] and the implementation in CASM bases mainly on these publications. There have been further improvements during the time up until now. This section is an abstract description of the cluster expansion method. It is in close conjunction to the recent work of Sanchez [77] who rigorously describes the mathematical foundations for a numerical implementation during his work from 1984 to date. The formal derivations begin with the definition of the underlying complete infinite basis set and defining constraints on it like completeness and normalizability. For the sake of space the interested reader is referred to the article by Sanchez [77] for a complete derivation. Only the main important results will be mentioned here as we start with the central equation for the cluster expansion

$$f(\vec{\sigma}) = \sum_{\alpha} j_{\alpha}(\mu) \phi_{\alpha}^{\mu}(\vec{\sigma}_{\alpha}), \quad (2.4.1)$$

which describes the expansion of a quantity in terms of effective cluster interactions (ECIs) $j_{\alpha}(\mu)$ and the corresponding cluster functions (CFs) $\phi_{\alpha}^{\mu}(\vec{\sigma}_{\alpha})$. Both ECIs and CFs are heavily impacted by the underlying symmetry of the structure one expands. In our case the quantity chosen to expand is the formation energy of the structure. With

$$j_{\alpha}(\mu) = \langle f, \phi_{\alpha}^{\mu} \rangle_{\mu} \quad (2.4.2)$$

we can compute the ECIs directly from a known formation energy that belongs to a CF. At this point the cluster expansion works only in a forward way. That means we have to input CFs ECIs and formation energies for a perfect (unrelaxed) structure and with this we can then

predict formation energies of yet unknown configurations. It serves as a theoretical model to prove the principle at this point only. For a practical use of the cluster expansion method we need to be able to input real formation energies, i.e. formation energies from relaxed structures. At this point a linear algebra formulation of the cluster expansion becomes necessary. Linear expressions can be inverted under specific conditions and when we can invert the principle equation we can feed formation energies from relaxed structures as input and predict formation energies of realistic systems. At this point we need to do some analytic mathematics to show we can actually transform the cluster expansion equations into a matrix formalism. For the full detail the interested reader is again referred to Sanchez [77]. In principle we need to show that the transformation

$$\langle f, g \rangle_\mu = \sum_{N_A=0}^N \mathcal{D}_N^\mu[x(\vec{\sigma})] (f, g)_x \quad (2.4.3)$$

is well defined and exists uniquely. With this existence we can transform the original equation

$$f(\vec{\sigma}) = \sum_{\alpha} j_{\alpha}(\mu) \phi_{\alpha}^{\mu}(\vec{\sigma}_{\alpha}) \quad (2.4.4)$$

into a vector space representation of the ECIs which reads

$$j_{\alpha}(\mu) = (\vec{f} \Lambda_N^{\mu} \Psi^{\mu} \Lambda_N^{\mu})_{\alpha}. \quad (2.4.5)$$

The sum of discrete enumeration index α has been transformed into matrix multiplication. The vector $\vec{f}$ represents a column vector of formation energies and the cluster functions have been algebraically transformed to $\Lambda_N^{\mu} \Psi^{\mu} \Lambda_N^{\mu}$ with $\mathbb{M}_N^{\mu} = \Lambda_N^{\mu} \Psi^{\mu} \Lambda_N^{\mu}$ where

$$\mathbb{M}_1^{\mu} = \frac{1}{\sqrt{2 \cos(\mu)}} \begin{pmatrix} \exp(\mu/2) & 0 \\ 0 & \exp(-\mu/2) \end{pmatrix} \begin{pmatrix} 1 & \varphi^{\mu}(1) \\ 1 & \varphi^{\mu}(-1) \end{pmatrix} \quad (2.4.6)$$

2.4.1. Inversion of CE

With the algebra done we can now formally write

$$\vec{E} = \mathbf{Z}\vec{V}, \quad (2.4.7)$$

which expresses the cluster expansion of the formation energy vector $\vec{E}$. This equation is now formally invertible to read

$$\vec{V} = \mathbf{Z}^+ \vec{E}. \quad (2.4.8)$$

With a constant offset $\vec{V}_0$ representing a fixed contribution to the cluster expansion, i.e. the energy contribution of a single point term and all non-constant contributions contained in $\mathbf{Z}_\phi^T$, this can be written as

$$\vec{V} = \vec{V}_0 + \mathbf{Z}_\phi^T \cdot \vec{Y}. \quad (2.4.9)$$

At this point we are formally able to invert the cluster expansion and input computed formation energies for relaxed structure that resemble real systems.

2.4.2. Cluster Approach to Statistical Mechanics

In the group of Anton van der Ven at University of California at Santa Barbara the Cluster Approach to Statistical Mechanics (CASM) code [78] is developed. The program package is divided into two main parts, one written in C++ which handles all computation and symmetry related tasks and the other part written in python which handles the C++ code inputs and outputs as well as interfaces to other python libraries and the queue system manager on a compute cluster. By default the code has an interface to VASP and I contributed an interface to FHI-aims with the same functionality. The CASM code implements the equations of the cluster expansion shown before. In addition to that the code also processes structures during generation of input and chooses only symmetrically distinct structures to be generated and computed. All relaxed structures are checked for symmetry equivalence again when processed as input for the cluster expansion. For some systems two distinct structures can relax into the same final structure of which only one is chosen as representation. The accuracy of the predicted formation energies depends on a number of parameters which are different for each structure and also depends on the species within the same structure. I.e. FCC aluminum has different parameters than FCC Copper. The CASM code provides the means to choose different basis sets. The user has to make the choice here as there is no working default for all problems. We used the CASM code [78] for the cluster expansion in chapter 5, which automates the construction and parameterization of effective Hamiltonians and implements these Hamiltonians in Monte Carlo simulations [79, 80, 81].

Training Data

The training data set is commonly the description of a collected set of DFT relaxed structures and their respective total energies. Most importantly all relevant accuracy settings have to apply to all structures the same way. This means that all basis and computational set-

tings have to be the same for all input data, i.e. same $\mathbf{k}$ -grid, same xc-functional and the same convergence settings like all forces below the same threshold. When these parameters are not consistently set the predictions of the cluster expansion will not be accurate in terms of the real system. There is no measure within the cluster expansion theory that would prevent errors in input data. The user has to make sure the input data is computed consistently.

Chemical Potential in CASM

Computation of the internal reference chemical potentials in CASM is done automatically. The values are taken from the DFT input data directly. Assume we have a structure with a single occupation degree of freedom, i.e. a surface with M sites that can be occupied or empty. Then the chemical potential of the occupying species is computed from

$$\mu^{\text{CE}} = \frac{1}{M} [E^{\text{DFT}}(\text{empty}) - E^{\text{DFT}}(\text{occ})] \quad (2.4.10)$$

In this case the chemical potential does contain only the DFT contributions. That means it is a zero-Kelvin computed value which does not contain any other contributions like zero-point vibrations or entropy which occur at finite temperature. If one wants to include these effects there are additional corrections needed as we will see later in chapter 5.

Basis Sets

The accuracy of predicted formation energies strongly depends on the fitting accuracy. For a given set of training data the fitting process depends on the basis set. The basis settings specifications are:

```
"basis_functions" : {
  "site_basis_functions" : "occupation"
```

```

},
"orbit_branch_specs" : {
  "2" : {"max_length" : 9.0},
  "3" : {"max_length" : 7.0},
  "4" : {"max_length" : 6.0}
}

```

for a common start. The choices the user has to make are the type of functions to use, key `site_basis_functions` with the choices `occupation` or `chebychev` that determine the accuracy to either the composition edges or the composition center. Further, the keyword `orbit_branch_specs` has more impact on accuracy. The first term, i.e. 3 designates the type of cluster, in this case a triplet while the term `max_length` specifies the maximum interaction range for this specific type. This way the user can include more or less different types of clusters and also modify the interaction range of those. These settings turned out to be most crucial to achieve a well represented data set and they can be very different for different types of structures (i.e. FCC, BCC and HCP) for even the same element. Also, for some structural problems only pair interactions are the dominant part (i.e. ionic crystals) while the particular problem in this work can only be well described when including up to quadruplet terms in the cluster expansion. In our case the choice of interaction and their respective lengths was made by trial and error to find the best possible representation of training data. Combinations of interactions of type 2, 3, 4, 5 and 6 were tested, also excluding some types, i.e. tying only 2, 3 and 5 with different lengths.

Exemplary Fitting generic Training Data

Let us consider a problem similar to what the cluster expansion will handle. The training data for this example was randomly chosen from an analytical function. We know there is a solution to the problem, i.e. a 'perfect fit'. This function and the training data points are shown in

figure 2.4.1. There are numerous different methods one can try and the goal of this exercise is not to find the best fit but rather to point out common pitfalls one has to keep in mind when using algorithms to parameterize arbitrary functions. We can see that a linear fit is, as

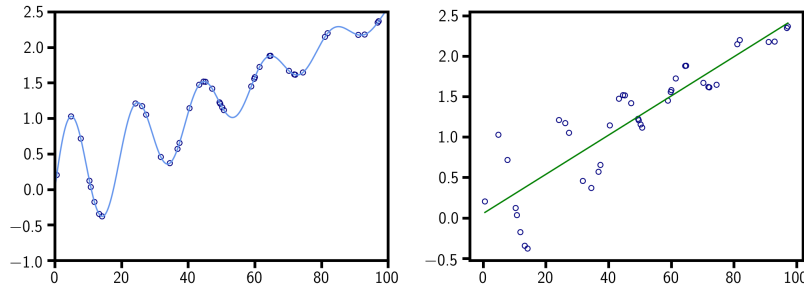


Figure 2.4.1.: Left: Analytical function to fit: $\exp(-x/50)\sin(x) + x/40$. Training data points from the analytical function are marked as circles, for the fitting process only these are used as input to the algorithm, **Right:** Linear fit capturing one term of the function only.

expected, just the average line. Besides not crossing many data points directly or passing within an acceptable distance, other data points are perfectly represented, i.e. the line touches the training data. Over all we can certainly do better by defining a more complicated function with more adjustable variables for fitting. Let's try a polynomial approach where we can increase the degree and have the algorithm figure out the coefficients a_n of each term in the form:

$$f(x) = \sum_n a_n x^n. \quad (2.4.11)$$

The results for degrees from 10 to 15 are shown in figure 2.4.2. The accuracy increases at first when increasing the number of variables,

i.e. the size of the underlying basis set. We see the same trend when looking at the plot with degree 12, accuracy has increased compared to 10 and 11. But a further increase now yield worse and worse results ending in a horrible overfit. The dangers lie in the testing procedure of the fitted data representation. When we look at plot with degree 15, we can see that although the training data is not well represented at all, there are three data points which are in direct crossing of the fitted lines, i.e. are perfectly represented. Should a testing procedure consist of asking for predictions at exactly these points only, the test would pass even with 95% of the data being very inaccurately represented. This example is to show that too few basis functions result in underfitting (degree 10) while too many basis functions result in overfitting (degree 15). The user has to find the right balance between these two. One main aspect that has not been discussed here is the choice of basis set function. While a polynomial approach might look like a good function to try there might be others which are suited better for this purpose. This will not be investigated further here as the principle problem of under- and overfitting will still remain.

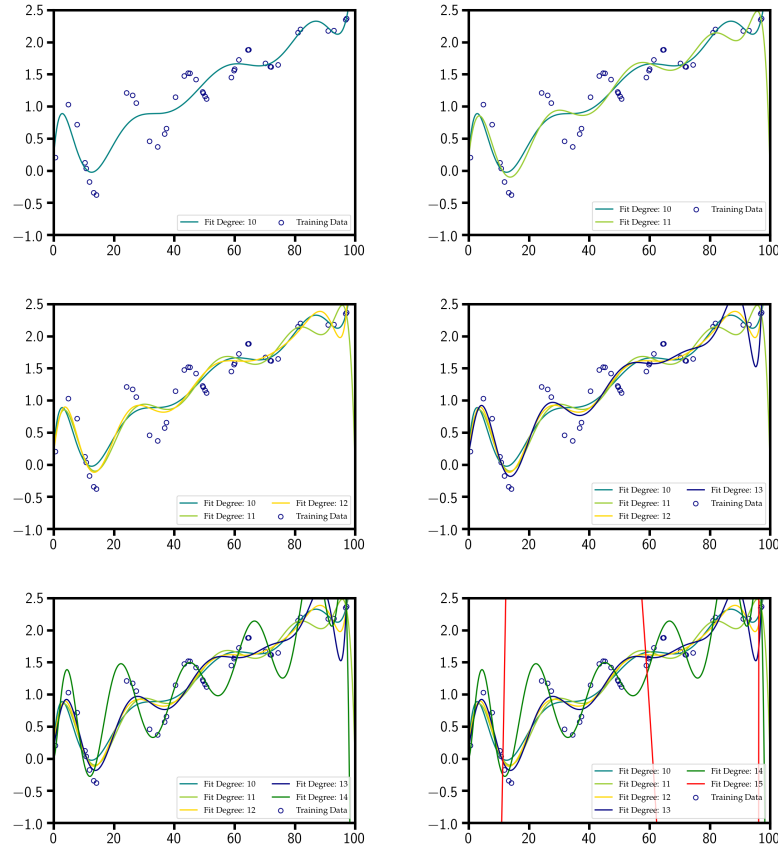


Figure 2.4.2.: Fitting example with the polynomial basis illustrating the problems of under- and overfitting. An increase of parameters of the basis set function does not necessarily result in an improvement of the possible representation. In general there is a sweet spot for each basis set beyond which one can not improve any further. The choice of a different basis set function might lead to better results but usually the information gained during fitting is not transferable between basis sets so the fitting procedure starts from scratch again.

A different model that is in a lot of peoples minds when talking about the use of algorithms to represent training data most accurately are neural networks. These are not analytically derived methods which are usually applied in this field, like Gaussian kernel regression [82]. Neural networks [83] are in general series of layers (nets) of neurons that can be active or inactive, much similar to how science explains the human brain. These networks are used in many applications to-day and trained models convince by their success. For completeness the results for a neural network trained on the same training data are shown here nevertheless in figure 2.4.3. The figure 2.4.3 aims to make

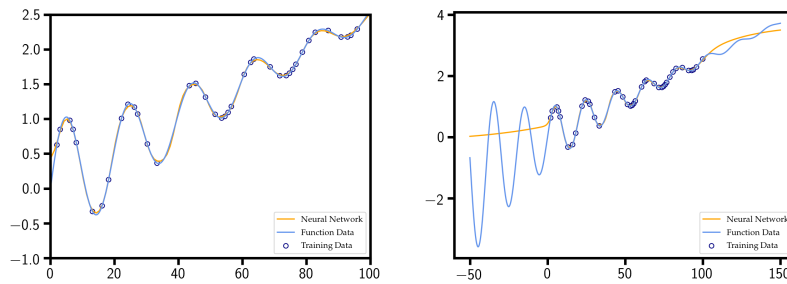


Figure 2.4.3.: Results from a neural network trained on the same data as the polynomial before

a general statement to all models representing data that was trained by any algorithm. Once one has achieved a fitting that represents the data well within the error margins that are sufficient for the user it can be used to make predictions out of the range it has been trained in, in this case outside of 0 to 100. The model does not have a barrier preventing this. The predictions will be inaccurate in many cases. What is even worse is the fact that in the range where I ask to model to make a prediction, there is no data to check against. No validation is possible and if left unchecked the predictions will quickly become simply absurd. This behavior can be seen in the right plot of figure

2.4.3. Predictions in the range $[-50,0)$ and $(100,150]$ are not appropriate compared to the almost 'perfect fit' in the range $[0,100]$. An implementation of a neural network fitting algorithm within the CASM package could be an efficient and effective contribution because the predictions of formation energies are always inside the boundaries defined by the input DFT data.

2.4.3. Model Selection

Now that we have seen the principle procedure behind the fitting mechanism we can talk about the way we choose to find the coefficients to the underlying basis functions. Let's imagine a function with n variables, called features. The routines that do the selection and fitting used in the CASM code are implemented in the `scikit-learn` package [84]. It is freely available under the GNU license. There are two algorithms mainly used in this work but there are more available in this package. The two used here are the genetic algorithm (GA) [85, 86] and the recursive feature elimination [87]. Both methods aim to find a function that represents the training data best but do it on different paths.

Genetic Algorithm (GA)

The genetic algorithm is initialized with a set of parents. Each parent is of a vector with length n and each entry specifies whether the feature with number n is set on or off. Initially this sequence is randomly chosen. So one parent is a list of zeros and ones of length n . In the first generation each parent is tested of its fitness, meaning how well does the parent represent the training data. With a comparison of the fitness of all parents the algorithm now generates children that inherit features from both parents and some of the fittest parent can also be retained. Additionally, there can be some mutation when children are generated. The worst ones are discarded. A ranking of the fitness

for purpose is performed after each evolutionary step employing the cross validation method introduced by van de Walle and Ceder [88]. In the next generation this process starts again checking the fitness of each participant, comparing and generating children while discarding the worst. This algorithm has no natural stopping criterion and the user provides the parameters for running it. The algorithm implementation in `scikit` is described in [85] and the algorithm itself stems from Ref. [86, 89].

Recursive Feature Elimination (RFE)

In contrast to the genetic algorithm this method starts with a function that has all features enabled. In the first elimination step the algorithm decides which of all features has the worst impact on the overall fitting accuracy and eliminates this feature. Now one does end up with a function of $n - 1$ features. The next step is the same as the first step, the worst feature is eliminated resulting in a $n - 2$ feature function. The RFE algorithm was initially developed for cancer research and is described in Ref. [87]. The user defines the final number of features at which stop.

Difference of GA and RFE

The main difference between these two algorithms is that for the RFE features are lost during the fitting process while in GA features are only turned off and can at a later stage be switched back on during the evolutionary algorithm if proven beneficial.

Cross Validation

A common approach to judge the accuracy of a trained model is the cross validation (CV). The testing of a model shall not be performed on the data it was trained on for obvious reasons. For this procedure one usually sets aside a part of the training data from the start to be

used later for validation. One has a training set and a test set in this approach where a ratio of 90:10 for training and testing is a good start. Then the model is trained on the training data part and after training has finished is validated on the test part to see how accurately the model can predict data we already know is correct. This is a common procedure and when using the CASM interfaces is automatically applied during training of the model. For GA algorithm the CV score validation is a part of the selection process during evaluation [88] while for the RFE algorithm it is computed once per feature selection and has no direct feedback within the algorithm.

2.4.4. Grand Canonical Monte-Carlo Simulations

The CASM code provides the ability to run Monte-Carlo simulations employing the Metropolis algorithm [90]. In this formalism the probability of a system to flip state is determined by a simple choice. Suppose we have a system containing two states only, A and B with their respective energies E_A and E_B with $E_A \neq E_B$. Suppose the system is in state A . Then, if $E_B < E_A$ the system will flip into state B , otherwise the probability of the system to flip is determined by

$$P \sim \exp\left(-\frac{E_A - E_B}{k_B T}\right), \quad (2.4.12)$$

with k_B as Boltzmann constant and T as temperature, as usual. Numerical implementations resemble the systems on grids containing each site as a grid point. In the simulation the flip probability is evaluated for each grid point individually based on the decision method presented above. Due to the nature of this decision there will always be a certain amount of fluctuation present during the simulation. One needs to introduce a method that can determine a convergence of the simulation within a certain set of parameters. This decision algorithm was developed by van de Walle and Asta [91]. For a more detailed mathematical and numerical discussion the reader is referred

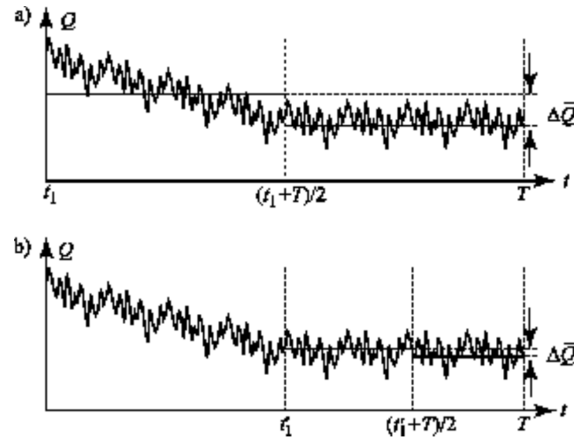


Figure 2.4.4.: Graphical interpretation of the convergence criteria for the Monte-Carlo Simulation, taken from van de Walle and Asta [91]. In upper simulation the blocks have significantly different mean values, rejecting the convergence while in the lower result the blocks have similar mean values, concluding that convergence has been reached.

to the original article by van de Walle and Asta [91]. This algorithm is used in CASM during Monte-Carlo simulations.

3. Surface states during CV measurements

The goal of this chapter is to shed light on the experimental measurements of cyclic voltammetry (CV) on IrO_2 and RuO_2 . The CV data

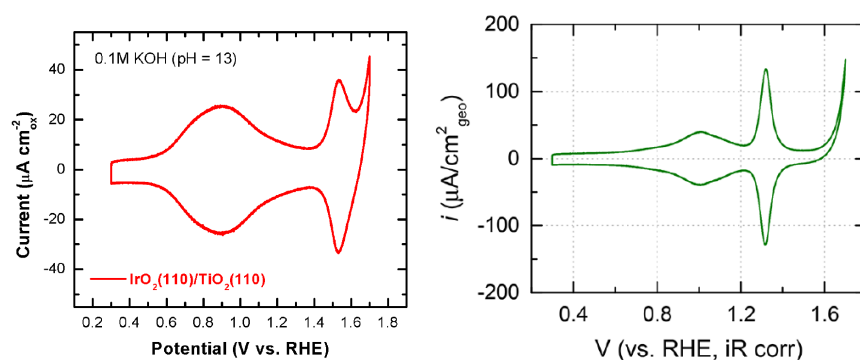


Figure 3.0.1.: Cyclic voltammogram of the IrO_2 (left) and RuO_2 (right) rutile (110) surface. The data is from our experimental collaborators, published in Ref. [15, 92].

were taken from a pristine surface sample that was grown by molecular beam epitaxy [24, 93] in the group of Schlom [94] and measured by Kuo [15, 92]. $\text{IrO}_2(110)$ films were grown by MBE on single-crystal $\text{TiO}_2(110)$ using a distilled ozone oxidant at a background pressure of 10^{-6} Torr. The flux of iridium was initially calibrated using a quartz crystal microbalance.

The epitaxial nature of the grown films was confirmed by in situ reflection high-energy electron diffraction (RHEED), low-energy electron diffraction (LEED) and X-ray diffraction [15]. The measurements show a fingerprint of the samples related to the activity on the surface before oxygen evolution starts. From experiments it is not possible to measure the hydrogen coverage directly during the CV so the exact atomic structure remains hidden. We will apply DFT and related computational techniques to reveal the nanoscale configurations of the phases involved during the CV measurement. Both CV measurements presented in this work are reversible.

3.1. Molecular Beam Epitaxy (MBE)

With this technique crystals are grown layer-by-layer in an ultra high vacuum environment. The growth itself is dominated by the bonding characteristics of the individual atoms or molecules involved in the process. The name combines the different processes involved during growth, the word beam refers to a constant kinetic stream of atoms or molecules that are assumed to have no interaction with each other [95] as opposed to a liquid flow of viscous nature. A flow of atoms or molecules is typically achieved by heating a source of the specific compound and this experimental technique has been developing since the late 1960s and is used for a wide variety of materials today. 3D materials can be grown with this method as was the case in the IrO₂ and RuO₂ samples presented in this work but also compounds of varying composition. Furthermore defects can be deliberately introduced that give rise to semiconductor research and more recently also quantum structures were engineered by MBE [96].

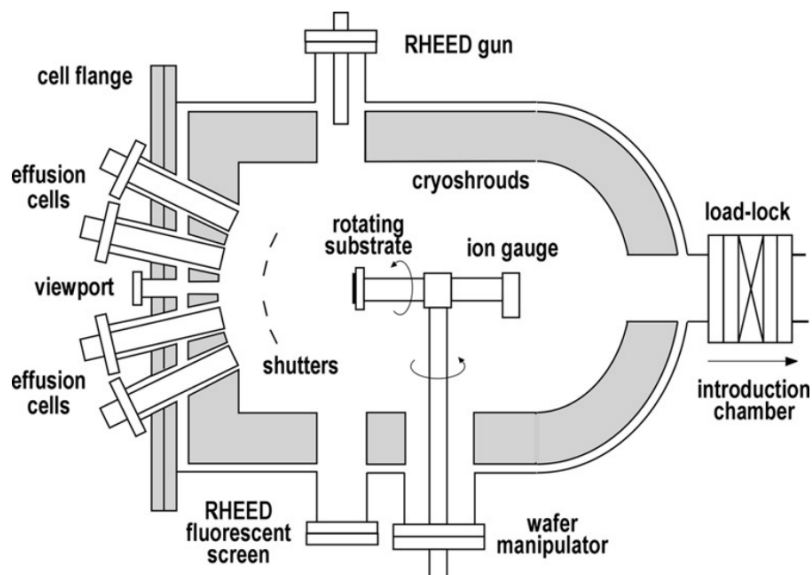


Figure 3.1.1.: Schematic view on an MBE growth chamber taken from [96]. Ions, atoms or molecules are introduced through the effusion cells and progress can be monitored via the RHEED screen or the viewport.

3.2. Cyclic Voltammetry (CV)

Cyclic voltammetry offers experimental insights into both the kinetics and the thermodynamics of an electrochemical system. Time dependence of the experimental data provides useful information about the kinetic processes. The shape of reversible voltammetric current responses is related to mass transport and diffusion in the solution phase. Since the CV is the result of transient and steady-state techniques the analysis of a CV is rather complex. During the initial phase of a CV measurement the diffusion layer moves into the solution as

depicted in 3.2.2. The thickness of this layer remains constant when the applied external potential is kept constant but moves with increase or decrease of the applied potential when the scan direction is reversed or scan speed is changed during the recording of a CV. Today's applications of CV include solid surfaces during electrocatalysis but have been applied to biological systems, polymers and different interface structures to name a few [27, 26]. Limits of the CV, in our case mostly the low resolution, need to be accompanied by other experimental techniques like spectroscopic analysis to gather information about the underlying structure of the catalyst. Simultaneous measurements of a CV and spectroscopy or diffraction can be used to gain further insight. Most common measurements are conducted in an electrochemical cell with three electrodes as exemplary shown in 3.2.1. These are named working electrode, counter electrode and reference electrode and are connected to a potentiostat. This setup allows to control the potential differences between the electrodes with minimal interference during the measurement.

3.2.1. Electrodes in the experiments

- Reference electrode: Reversible hydrogen electrode
- Counter electrode: Titanium wire or platinum gauze
- Working electrode: IrO_2 or RuO_2 catalyst grown by MBE

An electrical double layer at the working electrode occupied approximately 1 nm from the electrode into solution from the surface [27]. It consists of an inner layer called Helmholtz layer and an outer region, the Gouy-Chapman layer [26]. The exact extent to which this double layer reached into the solution depends on the concentration of ions in the electrolyte that influence its polarizability. The scan rate affects the thickness of this layer and at very low scan rates the natural convection begins to affect the shape of the experimental measurement.

Two laws determine the shape of a reversible voltammogram:

- Fick's Law:

$$\frac{d[A]}{dt} = C \frac{d^2[A]}{dx^2} \quad (3.2.1)$$

- Nernst Equation:

$$\frac{[A]_{x=0}}{[B]_{x=0}} = \exp\left(\frac{nF}{RT}(E(t) - E^0)\right) \quad (3.2.2)$$

Nernst described the surface concentrations of the two reactants A and B. In it, R is the ideal gas constant, T the temperature and n the number of electrons transferred per reaction step. F represents the Faraday constant. $E(t)$ is the applied potential and E^0 the reference potential. Fick's law is a diffusion equation and describes the mass transport towards the electrode with the coefficient C where x is the distance from the surface and t is time.

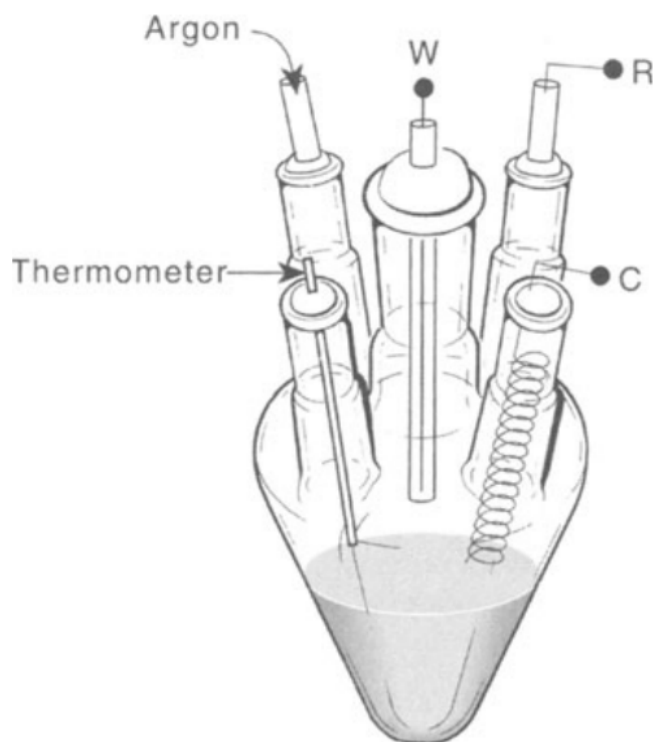


Figure 3.2.1.: Typical experimental bottle to perform cyclic voltammetry measurements with access ports for the electrodes [27]. **W**orking electrode, **R**eference electrode and **C**ounter electrode and marked. In this case argon is used as inert gas but it is of choice to the experimentalist.

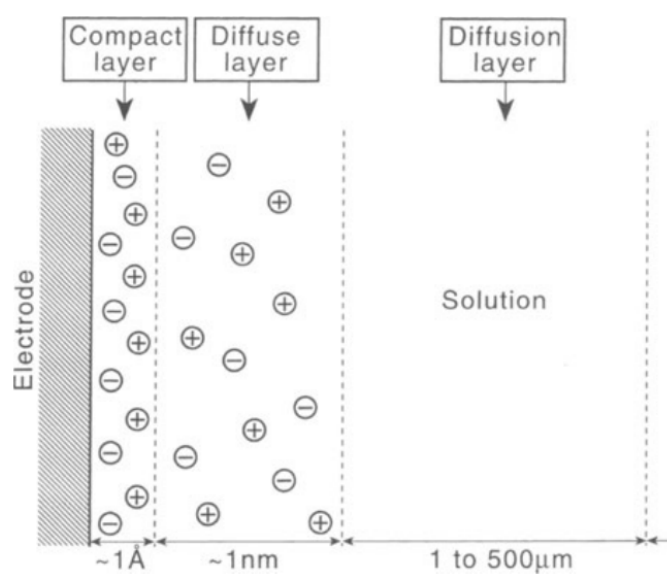


Figure 3.2.2.: Layers on the electrode. The compact layer has relatively high ordering and low spatial extension compared to the diffuse layer. Thicknesses are not constant but depend on the applied external potential. [27].

3.3. The influence of pH on the CV shape

The effect of pH can only be included implicitly at the theoretical level at this point. It is taken into account as a linear shift in the expression of the computational hydrogen electrode [97] but experiments have shown the effect to be more complicated. The resulting CV under different pH levels does depend in a nonlinear way on the level of pH but also has been shown to be influenced by the type of ions present in the electrolyte. Changing the pH to the same level with NaOH and KOH does also influence the peak positions on the resulting CV, see figure 3.3.1. In this work the effect of pH on the CV measurements will not be investigated further.

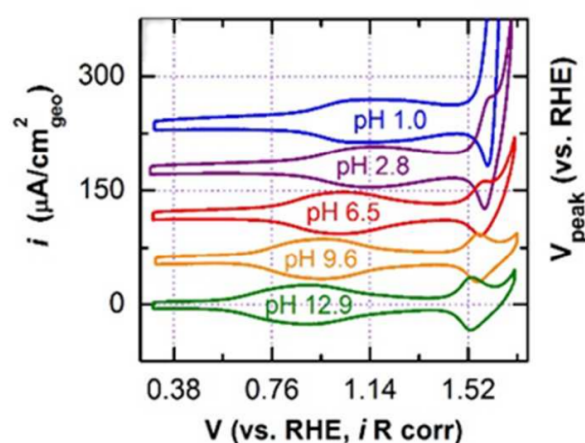


Figure 3.3.1.: Influence of pH on the shape of the CV, taken from [92]. With conditions becoming more acidic the broad peak shifts from about 0.9V towards 1.14V. While the onset of the oxygen evolution is not visible on the scale for the experiment in basic solution it is clearly visible in acidic conditions.

3.4. Interpretation of the CVs

In the left IrO₂ CV we see at about 0.9 V the center of a broad peak while at about 1.55 V we have a sharp narrow peak just before the OER starts at above 1.7 V. Comparing to the RuO₂ CV on the right we can make out distinct differences. The first broad peak has its maximum at about 0.1 V higher than on IrO₂ and the intensity is lower. From the lower intensity we can derive that the amount of electrons exchanged is lower on RuO₂ than on IrO₂. The second peak on RuO₂ has more than double the intensity than the peak on IrO₂ and occurs earlier at 1.3 V compared to 1.5 V on IrO₂. The oxygen evolution starts on both surfaces at about the same value of 1.7 V. From the similar shapes of the first peaks and the second peaks we can conclude similar processes on the surface to be involved during the CV measurement. However, the exact mechanism leading to the peak formation is unknown at this point. In general broad peaks relate to weaker interactions than sharper, more narrow peaks. In the following sections we are investigating these surface processes and want to derive the atomic configurations related to the positions and shapes of the experimental CVs.

3.5. Surface Phase Diagram Definition

One way to classify different surface structures of similar materials are phase diagrams. They are successfully employed in many different areas of materials science and surface science [31]. The mathematical definition of the surface energy reads

$$\Gamma^{\text{slab}} = \frac{1}{A} \left(E_{\text{slab}}^{\text{DFT}} - \sum_i n_i \mu_i \right), \quad (3.5.1)$$

wherein A denotes the surface area of the slab, $E_{\text{slab}}^{\text{DFT}}$ is the computed total energy of the slab while n_i and μ_i represent the number of constituent i and their chemical potential, respectively. The area of the slab is included to be able to compare different surface areas to each other and reference them to a given size. This approach follows along the lines outlined by Reuter and Scheffler [98]. Usually the energetics of the surface are plotted with respect to a reference. This reference surface is arbitrarily defined and depends on the application. Its surface energy Γ^{ref} is computed also by equation (3.5.1) and is regarded as constant during the construction of the diagram. The displayed quantity in the phase diagram then reads:

$$\Gamma^{\text{surf}} = \Gamma^{\text{slab}} - \Gamma^{\text{ref}}. \quad (3.5.2)$$

3.6. Chemical Potentials

The chemical potentials μ_i are obtained differently for each species. For some species like metals we can safely omit contributions of entropy and vibration in their bulk configurations. Other elements like hydrogen and oxygen are sensitive to these contributions because their reference states are in many cases gas phases where entropy and vibration are large contributions and can not be neglected.

Metals

For the phase diagrams presented in this work we take for the metals iridium and ruthenium the formula

$$\mu_M = E_{\text{atom}}^{\text{DFT}}, \quad (3.6.1)$$

wherein M represents Ir or Ru. We neglect other contributions like zero point energies and entropy. The chemical potential is thus the DFT total energy per atom.

Hydrogen

For the hydrogen chemical potential we take the formula

$$\mu_H = \frac{1}{2} \left(E_{H_2}^{\text{DFT}} + E_{H_2}^{\text{ZPE}} - TS_{H_2} \right) - U_{\text{RHE}}, \quad (3.6.2)$$

which besides the DFT computed energy of the H_2 molecule also includes terms for zero point vibrational energy $E_{H_2}^{\text{ZPE}}$ and entropy contributions ($-TS_{H_2}$). The vibrational energy is computed using VASP [99] and the gas phase entropy is taken from standard tables available from NIST [100, 101]. This chemical potential is related to the externally applied potential U_{RHE} via the computational hydrogen electrode [97]. This theorem assumes under standard conditions of pH=0, T=298K and $U=0$ the equilibrium



to hold. In this manner we can relate the hydrogen chemical potential to that of the electrons and by that to the external potential U_{RHE} since the electrons are driven by the external potential.

Oxygen

The oxygen chemical potential is now computed from the water molecule using the previously defined hydrogen chemical potential. We

use the water molecule at this point to avoid the known DFT overbinding error of the oxygen molecule computation. It then reads

$$\mu_O = E_{H_2O}^{\text{DFT}} + E_{H_2O}^{\text{ZPE}} - TS_{H_2O} - 2\mu_H. \quad (3.6.4)$$

Besides the DFT total molecular energy it also contains zero point energy and entropy contributions similar to the hydrogen chemical potential. The vibrational energy contribution is computed with VASP [99] as we did for the hydrogen molecule. The entropy contribution is derived from the gas phase by the ideal gas law. For this approximation we use the chemical potentials of gas phase and liquid water at 0.035 bar where they are equal at room temperature $T=298$ K [20]. The entropy is computed by

$$S_{H_2O} = (188.835 - 8.314 \cdot \ln 0.035) \cdot 1.036 \cdot 10^{-5}, \quad (3.6.5)$$

where 188.835 J/mol K is the gas phase entropy of water, 8.314 J/mol K is the ideal gas constant followed by the conversion factor $\ln 0.035$ from 0.035 bar to 1 bar from the ideal gas law. The factor $1.036 \cdot 10^{-5}$ resembles the conversion from J/mol K to eV. The reason for this conversion is the computation of the DFT water energy. This is done in vacuum for an isolated molecule in the gas phase rather than for water molecules in a periodic environment which would resemble the liquid phase where the total energy contributions from van-der-Waals and hydrogen bonding are difficult to compute [102].

3.7. Surface Structures

To model electrocatalytic processes we need suitable surface structures that represent the experimental setup as closely as possible. From experimental collaborators we know the basic parameters of the structure [15, 92]. It is IrO_2 and RuO_2 in rutile structure with (110) surface termination. The rutile structure is a tetragonal crystal system and is defined by two lattice parameters. For IrO_2 they are $a = 4.547$

Å and $c = 3.171$ Å. The atomic positions in fractional coordinates are listed in table 3.7.1. These differ from the values reported on the

x	y	z	Species
0.0	0.0	0.0	Ir
0.5	0.5	0.5	Ir
0.3095	0.3095	0.0	O
0.6997	0.6997	0.0	O
0.8123	0.1969	0.5060	O
0.1969	0.8123	0.5060	O

Table 3.7.1.: Atomic positions of the rutile IrO_2 tetragonal unit cell

materialsproject.org by a small amount ($\sim 0.01\text{\AA}$) due to us employing the RPBE [46] functional in this work. The lattice parameters show deviations of the same magnitude.

3.7.1. Surface Orientation

In general, orientations of planes in crystal structures are indicated by Miller indices [31]. The plane is constructed from the equation

$$p_{hkl} = h\vec{\mathbf{b}}_1 + k\vec{\mathbf{b}}_2 + l\vec{\mathbf{b}}_3, \quad (3.7.1)$$

where h , k and l are the Miller indices and $\vec{\mathbf{b}}_1$, $\vec{\mathbf{b}}_2$ and $\vec{\mathbf{b}}_3$ denote the reciprocal basis vectors of the lattice. The triplet (h, k, l) with $h, k, l \in \mathbb{N}$ then label a unique plane p_{hkl} in the structure. Surface orientations are classified by this unique triple (h, k, l) according to the plane they have been cut at. Exemplary, three planes of orientations (101), (110) and (111) are shown in figure 3.7.1. In practice we use scripts that take the unit cell and transfer it by linear algebra operations into the desired surface structure. This results in a 1×1 surface

of the desired orientation. To simulate the bulk structure below the surface the lower half of the slab will be fixed to bulk positions and not allowed to relax during simulation. Additionally, we introduce a hydrogen termination at the lower end to terminate the dangling bond which is left behind at the metal atom when cutting the slab. For this purpose all of the atoms in the 1×1 cell are fixed and only the terminating hydrogen atom is relaxed in z-direction. It is fixed in x- and y-directions to resemble the bond in the bulk system. Once this relaxation has reached convergence we have a 1×1 building block for the surface we want to model.

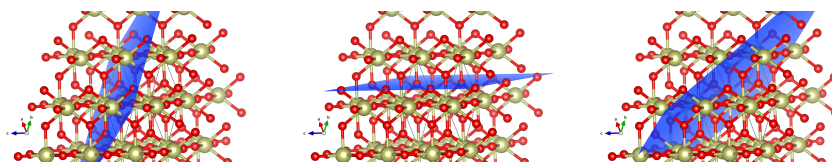


Figure 3.7.1.: Exemplary (h, k, l) -planes, **left** (101), **mid-**
dle (110), **right** (101).

From literature [67] we found that a layer of four O-Ir-O planes (a tri-layer) is regarded to be converged with respect to the surface adsorption energies. We use a denser k -grid corresponding to a $10 \times 10 \times 1$ grid for the 1×1 primitive cell for computations with FHI-aims. The computational cost for k -points is lower in FHI-aims than in VASP. For the force convergence criteria we generally set $0.01 \text{ eV/\AA}$. All structures are relaxed with first light and then tight [69] basis settings. For our simulations this surface building block is multiplied to yield the simulation surface we are using to compute the surface phase diagrams. We use this surface configuration previously as published by Rossmeisl [20, 97, 67]. The multiplication takes place in the

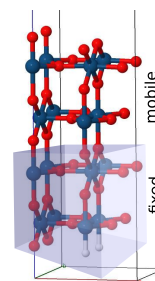


Figure 3.7.2.: The
final 1×2 cell for
computation

plane of the surface and effectively duplicates the primitive cell. We designate the result as 1×2 surface which corresponds to the multiplication factors used to generate it. The simulation cell for the rutile (110) surface is shown in figure 3.7.2.

3.7.2. Nomenclature for the surface configurations

Due to the amount of possible surface configurations we need a designation to distinguish them in a unique way. This is done here in a systematical way by following these rules:

- The surface consists of two identical halves, each has a coordinatively unsaturated site (cus) and a bridge site (b) shown in figure 3.7.3. Both halves will be separated by a dash and the designation A-B is equivalent to B-A where A and B name arbitrary adsorbates
- Considering only one half side now, the adsorbed water molecule has two possible designations. When it is adsorbed as a molecule on the cus it will be named OHHc. When it is deprotonated with OH adsorbed on the cus and H adsorbed on the bridge it will be designated OHcHb.
- Considering the complete 1×2 surface there are now three combinations for adsorbed water, namely OHHc-OHHc, OHcHb-OHHc and OHcHb-OHcHb. These are the three possible ground states of water on our 1×2 surface. They are shown in figure 3.8.1 and investigated for their total energies later.
- Other possible species on this surface during the CV measurement are OH and O. The oxygen atom stemming from the water molecule adsorbed on cus will not desorb at external potentials we consider here so the final species on each half side of the

simulated surface is Oc. The only other possible intermediate thus is OH which can occur in two configurations, as molecule OHc and deprotonated OcHb

- Considering one-electron steps (one-hydrogen desorption) we end up with many combinations of interest depending on how many rotational degrees of freedom are considered.
- Gibbs free energies of different surface configurations will be noted as shown in the following the examples:
 - For the clean surface: $G(* *)$
 - For adsorbed OHc-Oc: $G(*OH * O)$
 - For adsorbed OHcHb-OHc: $G(*OHcHb * OH)$

During electrocatalytic processes there are molecules adsorbing onto and desorbing from the surface without changing the underlying structure, i.e. no defects are being formed during electrocatalysis in our system. In our model we consider two coordinatively unsaturated sites that are mostly the functional sites but also two oxygen bridges which play a part. A picture of the simulation cell surface is shown in figure 3.7.3. On this surface the two cus are possible binding sites. Binding sites for hydrogen are the oxygens bound to the cus sites and the nearest bridge site. We do not consider hydrogen binding to the cus. At the simulated potentials there is no atomic hydrogen present, only water molecules are assumed to be present in the electrolyte.

3.8. The IrO₂ Surface

To simulate the cyclic voltammetry process we first need to understand the surface configurations when they are in contact with water. The calculation of the binding energy of water to the bare surface

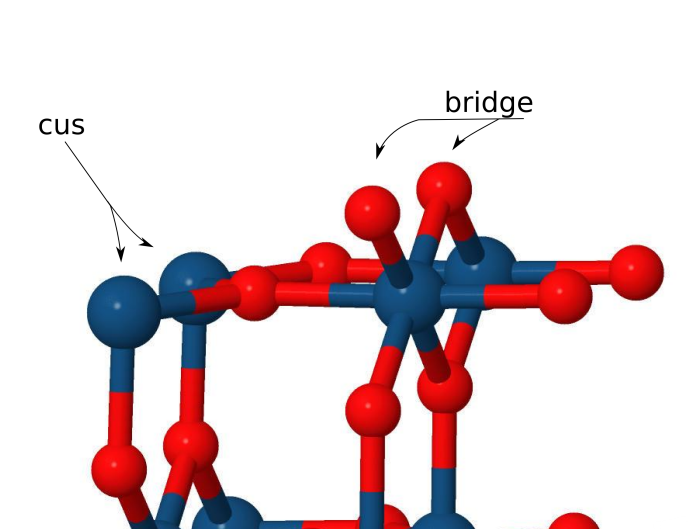


Figure 3.7.3.: Active sites *cus* and *bridge* on the 1×2 simulation surface. Oxygen adsorption takes place on the *cus* sites only while single hydrogen atoms can adsorb onto the *bridge*.

yields

$$\Delta G = G(*\text{OH}_2 * \text{OH}_2) - G(*) - 2 \cdot G(\text{H}_2\text{O}) \approx -2eV, \quad (3.8.1)$$

which indicates a strong favor for water molecules bound to both *cus* sites without applied external potential. Our starting structure therefore has three possible candidates shown in figure 3.8.1. The adsorption energies for all possible adsorbs are shown in table 3.8.2 later. Energetically these three structures are close, c.f. table 3.8.1. From a pure thermodynamical standpoint there is a single water ground state OHcHb-OHcHb . In experiment however, there will be a mixing of these states due to entropy at room temperature. The entropy

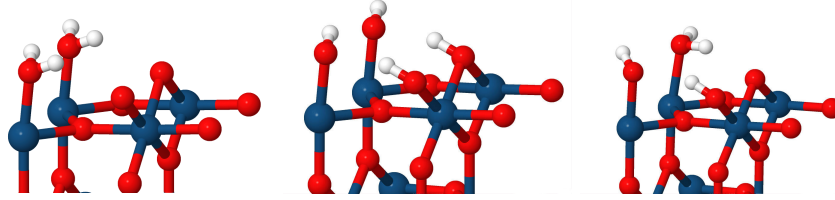


Figure 3.8.1.: The three possible adsorbates of water on the (110) rutile surface. **Left:** OHHc-OHHc, **Middle:** OHcHb-OHcHb, **Right:** OHcHb-OHHc.

introduced at room temperature is

$$k_B T = 8.617333 \cdot 10^{-5} \text{ eV/K} \cdot 300 \text{ K} = 0.025851999 \approx 26 \text{ meV}. \quad (3.8.2)$$

Structures within the range of 26 meV will be mixed statistically at room temperature. Considering these ground states we need to apply some corrections that are applied only for the adsorbed molecules. One is the zero point energy (E_{ZPE}) which we also have applied in the section defining the chemical potentials. In contrast to the definition of the chemical potentials we do not apply entropy corrections to adsorbed molecules. We assume the molecules to be fixed at their adsorbed position only able to vibrate. From VASP we can compute E_{ZPE} for adsorbed molecules similar as we have done for free molecules [99]. Our computed corrections for the three water ground states are listed in table 3.8.1. During the CV measurement hydrogen desorbs from the surface. This hydrogen was bound to the oxygen molecules of the water that adsorbed on the cus sites before the CV measurement started. We have a representation of this model in our three ground states listed in table 3.8.1 and figure 3.8.1. The next step is to apply external potential and see how the system reacts. The hydrogen chemical potential is linked to the external U_{RHE} by the computational hydrogen electrode introduced in equation (3.6.3). Applying this method we can now construct a phase diagram that depends

Type	E_{total} [eV]	ΔE [meV]	E_{ZPE} [eV]	ΔG [meV]
OHHc-OHHc	-8091220.404544575	0.005	1.351	7
OHcHb-OHHc	-8091220.284126629	120	1.347	123
OHcHb-OHcHb	-8091220.404601092	0	1.344	0

Table 3.8.1.: DFT energies for the three possible water ground states, their zero point energies E_{ZPE} and the Gibbs free energy difference ΔG to the reference state OHcHb-OHcHb.

on external potential U_{RHE} . Now we consider single steps in which one hydrogen atom desorbs from our surface model. We will thus include structures that have less and less adsorbed hydrogen on the surface. We built surface models of possible adsorbate combinations that arise from removing hydrogen from one of the configurations listed in table 3.8.1. The DFT total energies are computed and we apply E_{ZPE} corrections. A complete list of energies and corrections is shown in table 3.8.2. The final ingredient we need to construct a surface phase diagram is the number of constituents that define the reference surface OHcHb-OHcHb. The numbers of metal, oxygen and hydrogen atoms in conjunction with the total DFT energy and E_{ZPE} for the reference configuration define the reference Gibbs free energy for our model. We have chosen the reference surface to be the energetically lowest water structure OHcHb-OHcHb listed in table 3.8.1 and use the same chemical potentials as defined in section 3.6. With these we can now generate the surface diagram for the IrO_2 rutile (110) surface shown in figure 3.8.2.

Type	E_{total} [eV]	E_{ZPE} [eV]	ΔG meV
OHHc-OHHc	-8091220.404544575	1.351	0.005
OHcHb-OHHc	-8091220.284126629	1.347	120.5
OHcHb-OHcHb	-8091220.404601092	1.344	0
OHc-OHcHb	-8091203.151908742	1.009	0
OHc-OHHc	-8091202.959255768	1.012	192.7
OHc-OHc	-8091185.273844964	0.674	399.7
OHc-OHc_2	-8091185.673559043	0.674	0
Oc-OcHb	-8091167.811521313	0.454	20.6
Oc-OHc	-8091167.831150542	0.412	0
Oc-Oc	-8091150.327223435	0.151	0

Table 3.8.2.: DFT total energies and E_{ZPE} for all considered hydrogen configurations on IrO_2 . Structures OHc-OHc and OHc-OHc_2 differ in hydrogen orientation only. Equal hydrogen coverages are grouped and energy differences to the respective ground state are listed in the rightmost column.

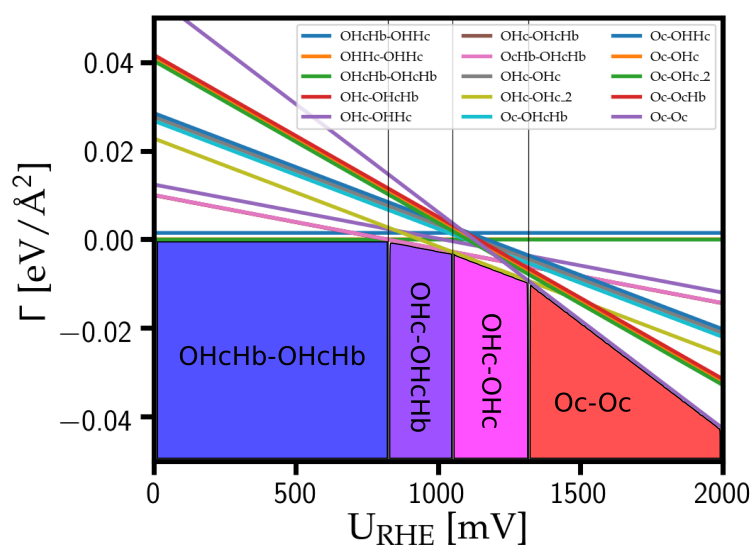


Figure 3.8.2.: Surface phase diagram of rutile IrO_2 (110) surface. Hydrogen coverage depending on external potential. Regions with different coverage are colored for visual distinction.

In figure 3.8.2 we see the hydrogen coverage depending on the externally applied potential. Thermodynamically there are four phases present on the rutile (110) of IrO_2 over the investigated potential range. They are shown in detail in figure 3.8.3. From 0 V to 0.85 V the surface

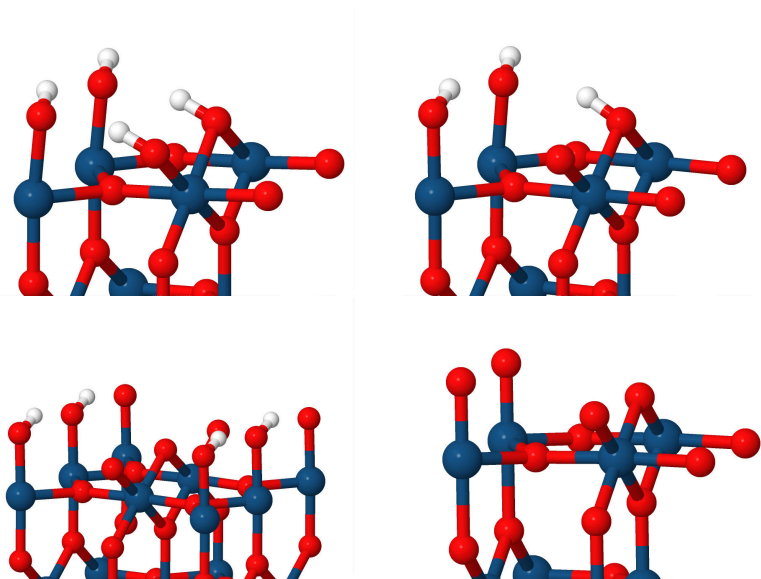


Figure 3.8.3.: Thermodynamically stable phases during the simulation of external potential, **top left:** OHcHb-OHcHb, **top right:** OHc-OHcHb, **bottom left:** OHc-OHc, **bottom right:** Oc-Oc surface configurations

remains in a water covered state. This phase is shown in figure 3.8.3 top left. At 0.85 V a phase change occurs, interpreted as the change of surface coverage from OHcHb-OHcHb to OHc-OHcHb. In the region from 0.85 V to 1.05 V the stable structure is shown in figure 3.8.3 top right. From 1.05 V to 1.3 V there is structure OHc-OHc stable shown in fig 3.8.3 bottom left. The last stable phase then is the oxygen remaining on the cus, label Oc-Oc figure 3.8.3 right. From the surface

phase diagram in figure 3.8.2 we can see that there are many competing phases present, some of which are energetically very close as we have seen before in the state analysis 3.8.2. A live system under experimental conditions will most likely not contain only the thermodynamically most stable structure at given external potential but rather a mixture of energetically close phases, especially when they merge near the potentials of phase change.

3.8.1. Comparing to the experimental CV

From our computations we can see the first phase transition occurring at 0.85 V and the second at 1.05 V. Both these values are lying inside the first broad peak suggesting a mixing of possible structure candidates. Our computed third phase transition at 1.3V lies between the first and the second peak. Moving from the transition energies of the peaks towards the shape of the peaks we have no way of deriving these from line crossings of a surface phase diagram. A possible solution to this is offered in chapter 5.

3.9. The RuO₂ Surface

The construction of the simulation cell follows the procedure described before for IrO₂. The parameters of the rutile unit cell used for RuO₂ are $a = 4.560 \text{ \AA}$ and $c = 3.134 \text{ \AA}$. The atomic positions in fractional coordinates are listed in table 3.9.1. These differ from the listed

x	y	z	Species
0.0	0.0	0.0	Ru
0.5	0.5	0.5	Ru
0.3059	0.3059	0.0	O
0.6941	0.6941	0.0	O
0.8053	0.1947	0.5008	O
0.1947	0.8053	0.5008	O

Table 3.9.1.: Atomic positions of the rutile RuO₂ tetragonal unit cell

values on `materialsproject.org` by a small amount ($\sim 0.01 \text{ \AA}$) due to us employing the RPBE [46] functional in this work. The lattice parameters show deviations of the same magnitude. We have seen the experimental CV for RuO₂ in figure 3.0.1 right at the beginning of this chapter. A structural analysis about possible surface states is similar to IrO₂ concerning the hydrogen configurations. The thermodynamic ground state will most likely be different due to the different metal in the host structure. Otherwise we proceed in the same footsteps we took before for the IrO₂ surface. First we compute all DFT energies for the different ground state candidates and list them in table 3.9.2. The thermodynamic ground state of water on RuO₂ is different from IrO₂. We chose the thermodynamic ground state OH_{Hc}-OH_{Hb} as our reference state. These ground states have been reported experimentally on RuO₂ by STM studies [103, 104]. The configurations of hydrogen on the surface have been generated as before

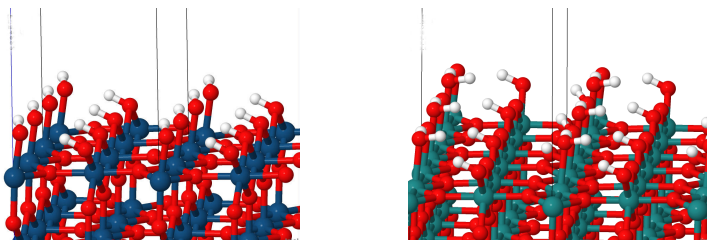


Figure 3.9.1.: Comparison of thermodynamic ground states of water, **Left:** OHcHb-OHcHb on IrO₂ and **Right:** OHHc-OHcHb on RuO₂.

for IrO₂ and their DFT energies and respective E_{ZPE} are listed in table 3.9.2. With these values we can assemble the surface phase diagram shown in figure 3.9.2. Similar to the IrO₂ surface, the RuO₂ surface shows a two-peak structure in the experimental CV. A major difference is that the broad peak is much less intense on RuO₂ than on IrO₂. The position of the first peak is at around 1.0 V. The broad peak is followed by a narrow peak at around 1.3 V. This peak height is much larger on RuO₂ than on IrO₂. The OER starts at about 1.7 V then which is similar to IrO₂. We can again interpret the phase crossings from our diagram and try to associate peak positions. From the surface phase diagrams there is a significant difference between RuO₂ and IrO₂. Comparing the phase diagrams we see that on RuO₂ we are left with only two phase transitions. Both fall energetically into the broad peak, the first transition at 0.85 V and the second at 1.1 V. It becomes clear that an accurate description of the second peak might not be possible with the applied methods. Again, due to the limit of the phase diagram we can only conclude to phase change energies but not towards peak shapes.

Type	E_{total} [eV]	E_{ZPE} [eV]	ΔG meV
OHHc-OHHc	-2063393.547955791	1.351	241.3
OHcHb-OHHc	-2063393.789251818	1.347	0
OHcHb-OHcHb	-2063393.520962101	1.344	268.3
OHc-OHcHb	-2063376.408446499	1.009	121.1
OHc-OHHc	-2063376.529596736	1.012	0
OHc-OHc	-2063358.614182792	0.674	371.2
OHc-OHc_2	-2063358.985406404	0.674	0
Oc-OcHb	-2063341.574138721	0.454	0
Oc-OHc	-2063341.470541579	0.412	103.6
Oc-Oc	-2063324.215946119	0.151	0

Table 3.9.2.: DFT energies and E_{ZPE} for all considered hydrogen configurations on RuO_2 . Structures OHc-OHc and OHc-OHc_2 differ in hydrogen orientation only. Equal hydrogen coverages are grouped and energy differences to the respective ground state are listed in the rightmost column.

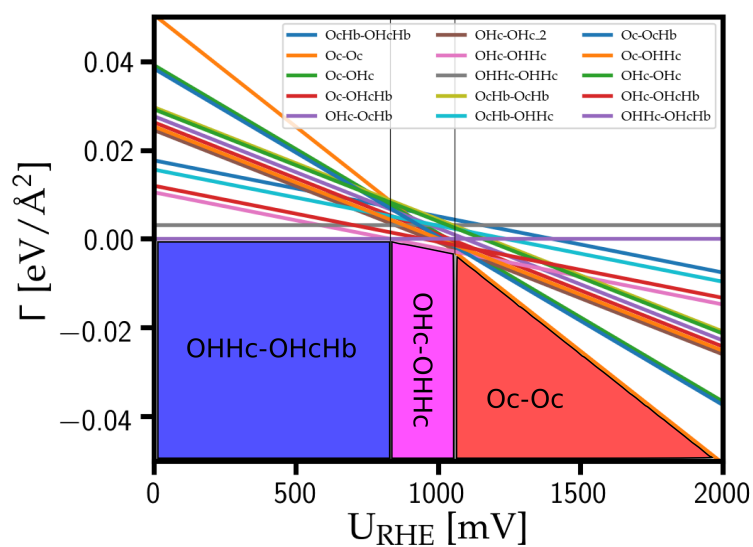


Figure 3.9.2.: Surface phase diagram of RuO_2

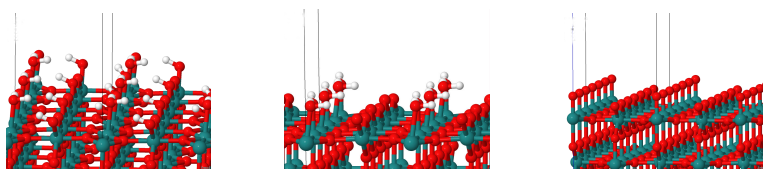


Figure 3.9.3.: The three thermodynamic stable states during the CV on RuO_2 . **Left:** OHHc-OHcHb, **Middle:** OHc-OHHc and **Right:** Oc-Oc. Note that the most stable water structure differs from the IrO_2 surface.

3.10. Adsorbed water on RuO₂

We have computed an STM image of a dissociated water molecule on the RuO₂ surface and compared it to the measurement presented in refs [103, 104]. There is a good agreement of the experimental data with our computation further confirming the thermodynamically stable ground state we have found by DFT. There was no such measurement available for the IrO₂ surface.

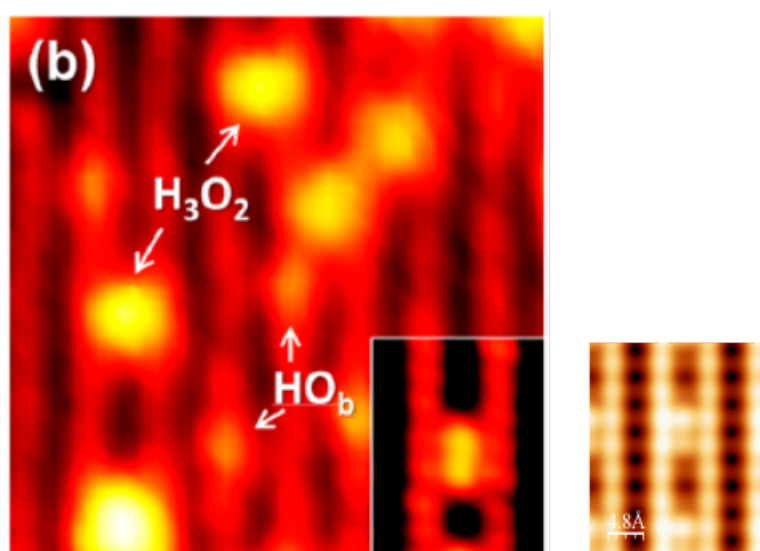


Figure 3.10.1.: Comparison of the experimental measurement of dissociated water molecules on the RuO₂ surface on the left and the computational STM on the right. The inset in the experimental data compared well to the computed STM image on the right.

3.11. Solvation energies

Although we are simulating an experimental process taking place in aqueous environment we do not model any additional water molecules beyond the adsorbates on the surface. The reason is that water structures are complicated [54, 55] and computational cost is rising strongly because of the possible degrees of freedom that water molecules present due to van-der-Waals binding. Hydrogen bonding and van-der-Waals effects give rise to a whole host of effects that are present in liquid water. The investigation of water layers above the surface is not subject of this study. Computationally one can not handle all the degrees of freedom that arise from a surface that is covered by water molecules explicitly. Traditionally one simply ignores this water layer and computes the slab structures in vacuum as is done by other groups [105, 106, 107] and as we have done in this work up to this point. Recently, there has been a lot of development of implicit solvent models. These models all stem from the same physical idea. The water molecules are not modeled explicitly but instead one assumes they behave like a constant background with a dielectric constant. Based on this assumption various models have been developed. We have tested a readily available implicit solvation model [59, 60] for VASP with our simulation surfaces of IrO₂ and RuO₂ to investigate the effects. This model simulates a constant background and has as parameter a dielectric constant. It was set to 78.4, the dielectric constant of water. More sophisticated models are implemented in the jDFTx code [108]. They are mentioned here for completeness.

- LinearPCM [61, 62, 63]
- Nonlinear PCM [61, 64]
- SaLSA [64]
- ClassicalDFT [65, 66, 65]

The resulting solvation energies are shown in figure 3.11.1. While ad-

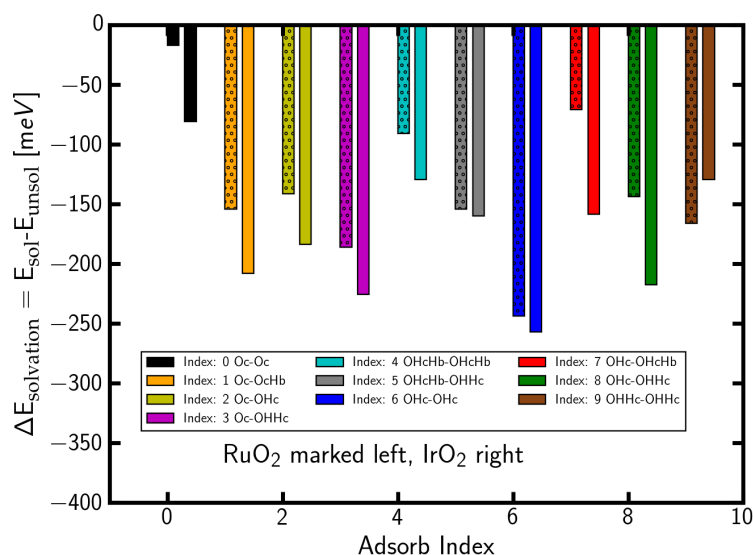


Figure 3.11.1.: Computed solvation energies from VASP-sol for IrO₂ and RuO₂ adsorbate configurations. The general trend is a lowering of total energy and while there are differences in solvation energy of up to 50 meV when comparing the same adsorbate on IrO₂ and RuO₂ these are not significant enough to change the energetical ordering of surface structures in the phase diagrams for either surface.

sorbates like OHcHb-OHHc and OHc-OHc do not show strong energy differences between IrO₂ and RuO₂ there are in general differences of ~ 50 meV when comparing different metal surfaces. The largest difference of ~ 100 meV occurs for OHc-OHcHb. This is most likely due to different charge transfer from the adsorbates to the surface metals. For all adsorbates IrO₂ present the lowest solvation energies, only for

the OHHc-OHHc adsorbate we see the situation reversed. The phase diagrams presented in this chapter do not change significantly when including solvation energies. In particular the ordering of phase transitions is unaffected. For this work we do not consider solvation energies any further.

3.12. Results

We have presented surface phase diagrams for IrO_2 and RuO_2 in similar fashion. Comparisons to experiment show good agreements with the first peaks of RuO_2 and IrO_2 , respectively. The second peak can not be assigned. Perhaps a different DFT functional could help capture these better, other methods that go beyond DFT could also help describing these transitions more accurately. From these diagrams we assign the underlying mechanism and the atomic structures resulting in the formation of the peaks during the CV measurement. The respective surface coverage is labeled in the phase diagrams and shown color coded in the experimental CVs. The inclusion of solvation energies does not explain the deviations we observe for the positions of the phase crossings for both materials when comparing to the experimental CVs. Making a direct comparison of the experimental CV to the computed phase transitions we arrive at figure 3.12.1 (top) for IrO_2 and figure 3.12.1 (bottom) for RuO_2 . Here we see the discrepancy between the experimental and theoretical results clearly. A first remark needs to be made about the number of phase transitions. While we start with a water covered surface for both metals we do not observe four phase transitions in either case. IrO_2 shows three transitions while RuO_2 shows only two. From a thermodynamical point of view this seems surprising since we are assuming one-electron processes. Especially the particular construction of our surface configurations is tailored that way. Some phase transitions do not appear stable within this description. From the electrochemical view it means that there are processes involved that DFT does not capture completely and the theoretical description can be improved. As a similarity between both surface diagrams we can remark that in the occurring phase transitions it is the bridge-bound hydrogen that is desorbed first. In RuO_2 the water covered surface is OHHc-OHcHb transforming to OHc-OHHc at 0.85 V. On IrO_2 it is first OHcHb-OHcHb transforming to OHc-OHcHb at 0.85 V which trans-

forms to OHc-OHc at 1.05 V. In both phase diagrams the hydrogens adsorbed to the oxygen on cus are described in a single process. On RuO₂ three hydrogen desorb in the step at 1.1 V and on IrO₂ two hydrogen at 1.3 V. The second remark is that although we are investigating configurations that are structurally very similar in configuration the transition metal influence is clearly in play. We have in both cases different thermodynamic ground states and we also have a different transition state. The number of phase transitions is also different suggesting different electronic influence of the metals near the surface. We also conclude from the phase diagrams that for reactions that occur at potentials above 1.3 V for IrO₂ and 1.1 V for RuO₂ the surface is covered by oxygen atoms only.

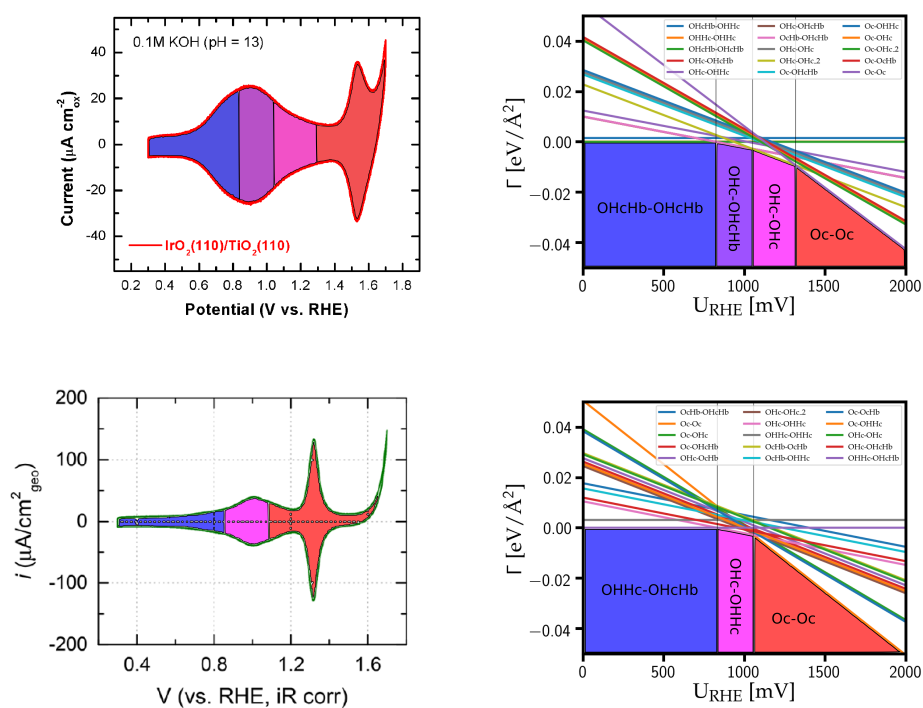


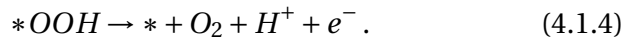
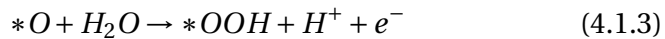
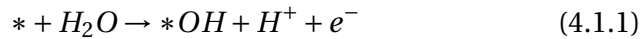
Figure 3.12.1.: Direct comparison of experimental CV and computed phase transitions on IrO_2 (top) and RuO_2 (bottom)

4. Ti-IrO₂ alloy

This chapter will focus on the oxygen evolution reaction. It is a theoretical analysis which is validated by comparison with the experimental data as done in the chapters before. VASP [33, 34, 72, 73] is applied for all computations of DFT total energies here. As before, the RPBE functional [46] is used with 400 eV plane wave cutoff energy. Reciprocal space sampling is equivalent to a $8 \times 8 \times 1$ sampling of the primitive cell and we make use of the dipole correction [68] to avoid slab-slab interactions.

4.1. Oxygen Evolution Reaction

The theoretical description of the OER was introduced by Nørskov [97] for pure metals. Due to its success when comparing theoretical results to experimental findings the application of this theory was quickly widened to many other structures apart from pure metals, like metal oxides [109, 110], metal hydroxides [111, 112] or perovskites [113]. For some classes the theory worked well [20] while it failed to explain the experimental measurements for others [113]. The equations describing the OER are



The DFT energy of liquid water is computed by the recipe presented by Man [20]. The H^+ chemical potential is computed consistently in the same manner as we have done earlier in equation (3.6.2). Corrections for zero-point energies to the adsorbs are for $*O$: 0.05 eV, $*OH$: 0.35 eV and $*OOH$: 0.4 eV. These are consistent with the corrections that were applied when adding the zero point energies within the cluster expansion framework in chapter 5.

4.2. Pure IrO₂

IrO₂ is regarded as standard reference for OER activity [15]. From the work presented in the previous chapters we know the surface configuration of IrO₂ at the experimentally determined OER potential. The catalytically inactive cus sites are saturated with oxygen atoms such that the OER cycle described in equations (4.1.1)-(4.1.4) takes place on the active cus site. This surface model is consistent with other previous work and the computations done here follow the same framework [20, 109, 110]. The images shown in figure 4.2.1 show the

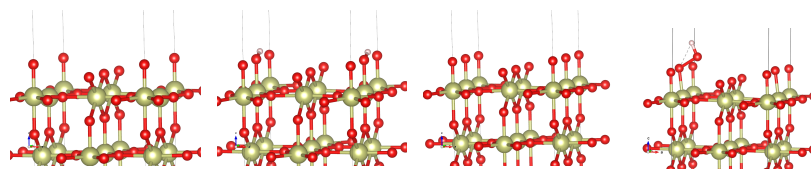


Figure 4.2.1.: OER steps to compute energies for equations (4.1.1)-(4.1.4) on the pure IrO₂ rutile (110) face.

configurations used to compute the adsorption energies for single reaction steps. The pictures present the single reaction steps for equations (4.1.1)-(4.1.4). First we start with the active cus site being empty. Then a water from solution is split, the remaining OH is adsorbed onto the active cus site forming $*OH$ while the proton travels through the

electrolyte onto the cathode surface. The second step then splits off the hydrogen from the just adsorbed $*OH$ forming $*O$ which also travels to the cathode, recombining with a present hydrogen to H_2 or adsorbing and recombining later. The third step again needs a water from solution to be split and the OH adsorbed to the $*O$ on the active cus site forming $*OOH$ in the rate limiting step for the reaction on this specific metal surface. The left over hydrogen behaves similar to the ones split off before. In the last step the $*OOH$ desorbes from the active cus site and is stripped off its hydrogen. The molecular oxygen travels through the electrolyte into the gas phase while the hydrogen again recombines with present adsorbed hydrogen at the cathode or adsorbs on it for later recombination. The empty active cus site is left and the cycle can start anew. During this process a total of 4 electrons are exchanged and the overall thermodynamic potential to overcome is 4.92 V, the equal of four single 1.23 V steps.

4.2.1. Kinetic considerations

Catalytic reaction on surfaces are dominated by the kinetics of the reaction. While we are not computing kinetic barriers from DFT to calculate the theoretical overpotentials for our metal surfaces we assume the Brønsted – Evans – Polanyi (BEP) [18] scaling relation holds. It states that the kinetic barrier for a reaction scales with the adsorption energies. Kinetic barriers can be computed from DFT but they need much more computational time than only adsorption energies. Assuming the BEP relation holds for our case we can therefore derive overpotentials from adsorption energies. Comparing the results from these computations to experimental values we see that the trends are covered correctly. Materials with low theoretical overpotentials show higher activities experimentally than materials with higher theoretical overpotentials. With these considerations a special note has to be made about the structure for the $*OOH$ adsorb. From a pure thermodynamic point of view the structure shown in figure 4.2.1 right is

not the one of lowest energy. The absolute lowest energy structure is shown in figure 4.2.4 and shows a hydrogen transfer from the $\ast\text{OOH}$ adsorbate to the bridge oxygen forming $\ast\text{OO} + \ast\text{H}$. These structures have been investigated under kinetic considerations by Dickens [114]. Their work finds that the linear scaling relation only applies well when the *up* configuration for the $\ast\text{OOH}$ adsorbate is used, other configurations do not agree well in the linear limit, as shown in figure 4.2.2. For this reason we will use this specific $\ast\text{OOH}$ configuration in our computations for theoretical overpotentials.

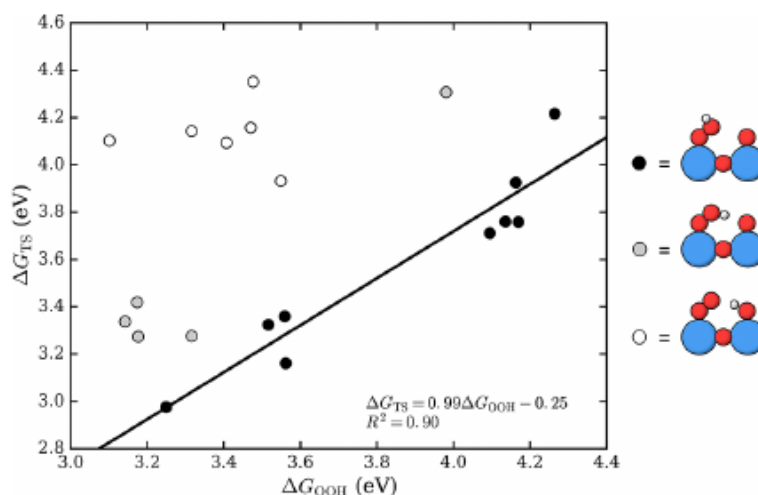


Figure 4.2.2.: Scaling relation for different adsorbate configurations of $\ast\text{OOH}$, taken from Ref. [114]. The linear scaling only applies to the $\ast\text{OOH}$ adsorbate in *up* position, others show large deviations and are not taken into account for further analysis in this work.

4.2.2. Theoretical overpotential

The theoretical overpotential η_{th} is computed from single Gibbs adsorption energy steps. We know from literature [12] that the value of energy required to split a water molecule is 1.23 V per electron. Since the OER equations describe four single-electron processes the perfect catalyst would have a reaction step energy of 1.23 V per step. This perfect catalyst does not exist and real materials show different Gibbs energies needed for each step. The reaction step showing the largest energy is called the rate-limiting step. For the pure IrO₂ and Ti-alloyed IrO₂ this step is the adsorption of *OH* onto an existing **O* on the surface to form **OOH* which is described in equation (4.1.3). The energy needed for the OER cycle to return to its initial state is fixed to 4.92 V because of the four electron process. The single Gibbs free energy steps are computed from equations (4.1.1)-(4.1.4) and are commonly displayed in plots similar to figure 4.2.3. The theoretical overpotential is computed from the rate limiting step by subtracting the one-electron split-energy of 1.23 V from the step energy:

$$\eta_{\text{th}} = \max[dG_{OH}, dG_O, dG_{OOH}, dG_{\otimes}] - 1.23V. \quad (4.2.1)$$

The reported values for the theoretical overpotential of IrO₂ are 0.5 V as is shown in the volcano plot in figure 4.3.1. This overpotential is computed using the deprotonated hydrogen adsorbate configuration of the **OOH* adsorbate shown in figure 4.2.4. In light of the previous discussion about the BEP relation and the consistency of adsorbates with the linear scaling relation discussed in section 4.2.1 we chose to not use the deprotonated **OOH* configuration but instead take the **OOH* in its *up* configuration. This is why we do not arrive at the reported value of 0.5 V for the theoretical overpotential but have a value of 0.83 V for pure IrO₂ at this point.

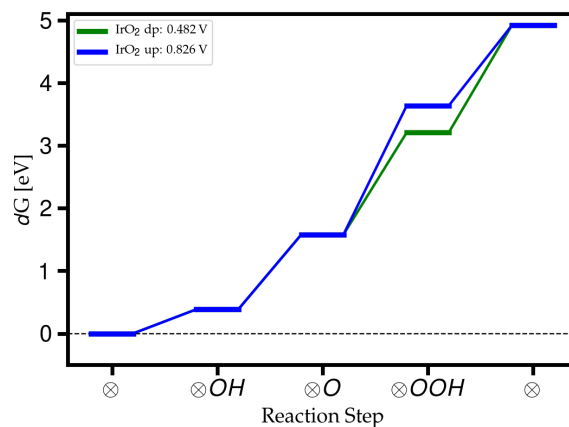


Figure 4.2.3.: Gibbs free energy reaction steps on IrO₂ for the two different adsorbate configurations from figure 4.2.4. The deprotonated configuration is labeled *dp*. The theoretical overpotentials are given in the labels.

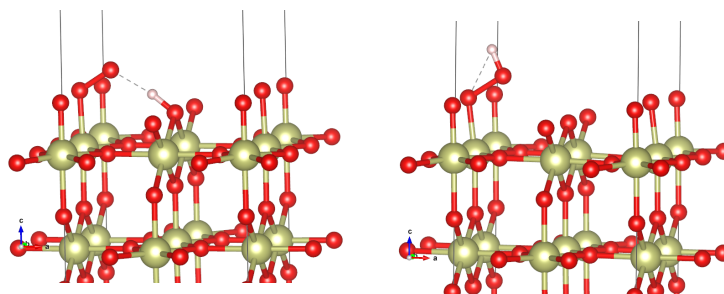


Figure 4.2.4.: Comparison of adsorbates on pure IrO₂, left: the deprotonated and right: the *up* configurations. The right one is used for overpotential computations in this work.

4.3. Can we improve IrO₂ by alloying Ti?

4.3.1. Volcano plot

A typical volcano plot is shown in figure 4.3.1 from Man [20] as it was introduced in chapter 1. The x-axis descriptor is usually chosen to be $\Delta G_{*O} - \Delta G_{*OH}$ and the y-axis is the negative overpotential $-\eta_{th}$. It is generally used as a classification of activity for investigated materials. It gives a good overview of the materials performance. To the top of the volcano there are the most active materials and along the sides performance drops. The materials investigated so far in this work are IrO₂ and RuO₂. Both are located near the top of the volcano and are reported as the most active OER catalysts in literature as well. Another material to be noted at this point and will become of interest is TiO₂. Located on the right side of the volcano at about 1.2V overpotential it is generally considered inactive with respect to the OER. It has to be noted that there are many different structures compared in this volcano plot and they are not all rutile or not necessarily stable under OER conditions. For a complete list of structures and metals we refer to Man [20]. In this work we focus not only on the lowest theoretical overpotential we can achieve by computation but also the experimental viability with respect to stability and non-toxicity. From this volcano plot shown in figure 4.3.1 we can see that IrO₂ is slightly to the left side of the top of the cusp. If it was possible to 'push' the binding energies on IrO₂ a little bit further towards the

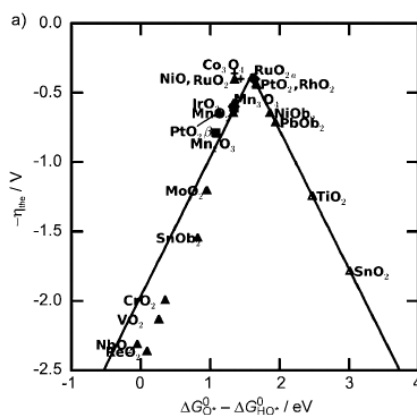


Figure 4.3.1.: Volcano plot for different metal oxides from Man [20].

right one could end up with a material presenting lower overpotential and thus higher activity. The choice made by the experimentalists fell to titanium because it does not jeopardize the acidic stability of IrO₂. It is also non-toxic. The idea is to stabilize the adsorbate of a specific reaction step and thereby lowering the overpotential without influencing the other reaction step energies negatively. Looking at the reaction steps the only adsorbate that has large enough spatial extension is *OOH. It reaches over towards the bridge oxygen which enables interaction with the underlying atoms. The picture of this adsorbate is shown in figure 4.2.1 right.

4.4. Experimental Data

From the experimental results presented in figure 4.4.1 we can clearly see an improvement in activity achieved by mixing the IrO₂ surface with small amounts of titanium. The optimal amount of Ti is about 14% to achieve the best results. The fabrication procedure is identical to the IrO₂ host material grown by molecular beam epitaxy. The only difference here is the replacement of a small portion of iridium by titanium during growth. The surface termination, reaction path and active site are assumed to be identical among the pure IrO₂ and the alloy surface structures.

4.5. Possible motifs for Ti-enhanced IrO₂

To investigate possible positions of Ti in the IrO₂ host structure DFT total energies of random replacements of Ir with Ti were computed. The total energy difference per atom to the lowest energy for each structure is listed in table 4.5.1. Structure p1 has the overall lowest total energy per atom and is used as the reference here. A complete set of pictures of all structures used to compute the total energies in

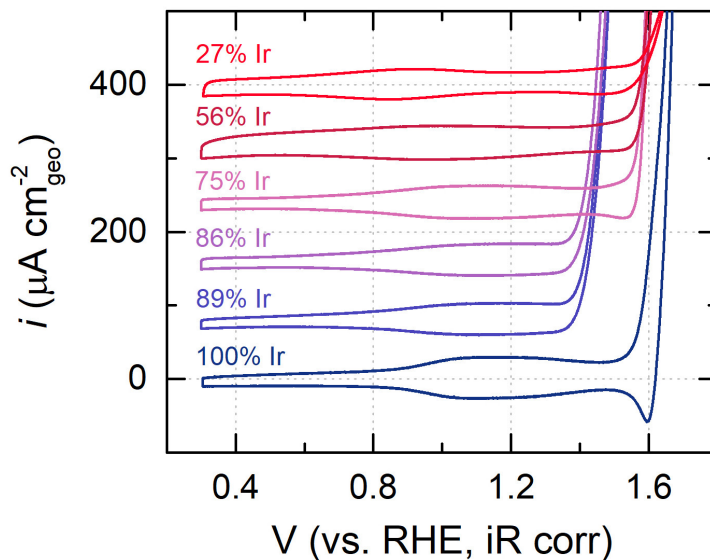


Figure 4.4.1.: Experimental CV of Ti-alloyed IrO₂. The improvement in activity is clearly visible with its maximum at 14% titanium content. Measurements taken by the group of Jin Suntivich.

table 4.5.1 is shown in the appendix V. From the small energy differences of all configurations presented in table 4.5.1 we conclude that titanium has no preferred position in the host structure and will be randomly distributed resulting in a statistical mix. To relate theoretical DFT data to experimental measurements we need a motif that can represent the modified surface of the Ti-alloyed IrO₂ and that allows for the computation of adsorption energies for the OER reaction steps in equations (4.1.1)-(4.1.4). From these energy steps the theoretical overpotential can be computed in the same way as for the pure IrO₂ host structure. There are only two sites on this surface that are directly influential to the OER, one cus and one bridge site.

The motif here is to replace one Ir in each site by a Ti atom. This results in two configurations representing the new alloy surface shown in figure 4.5.1. This motif has been used before by Halck [110] to represent nickel and cobalt replacements in a RuO₂ host structure.

Name	ΔE_{DFT} [meV / atom]
a	3.542
b	5.542
c	7.851
1	8.397
2	6.286
3	6.242
p1	0.0
p2	3.710
p3	2.763

Table 4.5.1.: Total DFT energy per atom for the computed structures shown in appendix V.

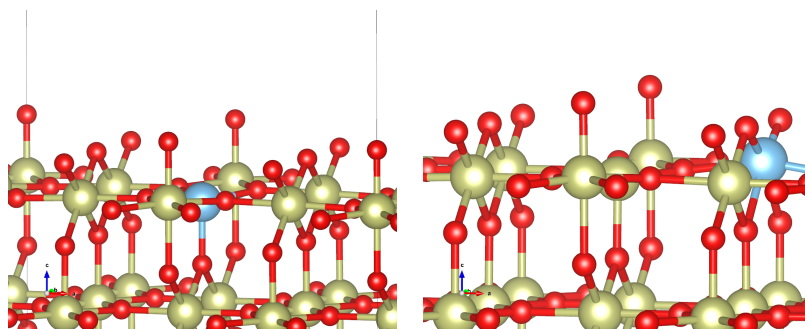


Figure 4.5.1.: Surface model of the Ti-mixed IrO₂ host structure, **left:** Ti in cus, **right:** Ti in bridge position

4.5.1. Theoretical overpotential

Employing the same DFT techniques as for the pure IrO₂ structure we compute the Gibbs free energy steps for the OER on Ti-mixed IrO₂ shown in figure 4.5.2. We can see the effect of the titanium in bridge position lowering the adsorption energy for the **OOH* reaction step. This effect positively influences the OER activity. The theoretical overpotential of pure TiO₂ is about 1.2 V as seen in the volcano plot in figure 4.3.1. This is regarded as inactive at the overpotential for the OER on pure IrO₂. Assuming a statistical mixing of titanium in random positions there will be some titanium in the cus site and some in the bridge site. The total amount depends on the amount of titanium mixed in during MBE growth. The general assumption is that there will be an even number of cus and bridge sites present. While the cus sites inactivate the material the bridge sites improve the activity due to the stabilization effect of the **OOH* adsorb. At a critical percentage the amount of improved cus sites by stabilization will be at an optimal value and the activity is thus maximal. This percentage shows when about 14% of titanium is mixed in which we can see from figure 4.4.1. Lower values are driving the activity towards that of pure IrO₂ while increasing the amount of titanium will ultimately result in the TiO₂ surface which we already know is inactive. We also consider the configuration of titanium in cus site as inactive because the binding of oxygen to titanium is unfavorable as we see in the volcano plot.

To be consistent with our previous computations we have also chosen the *up* configuration of the **OOH* adsorbate on the doped surface as shown in 4.5.3. Comparing both materials by their theoretical overpotentials we can explain the experimental findings. The theoretical overpotential of pure IrO₂

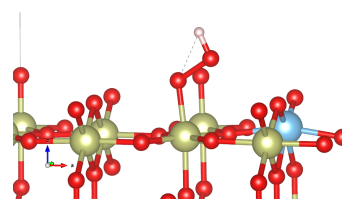


Figure 4.5.3.: **OOH* state on Ti-mixed IrO₂ on rutile (110)

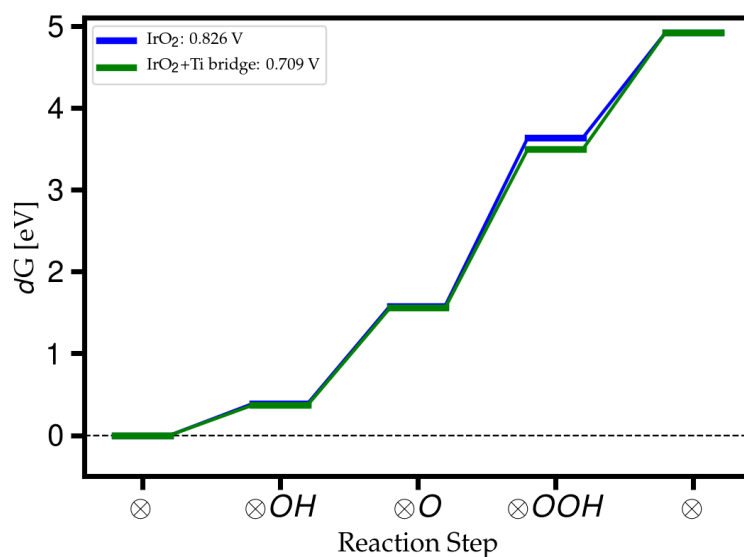


Figure 4.5.2.: Gibbs free energy reaction steps on Ti-mixed IrO₂ in bridge position. The stabilizing effect resulting in improvement of the OER activity influencing the *OOH adsorbate is visible. The energy is lowered by 117 meV for the *OOH adsorb.

$\eta = 0.83V$ is lowered by 117 meV by the stabilizing effect of the Ti atom in bridge position. This value of 117 meV corresponds to the experimentally observed reduction of the OER onset presented in figure 4.4.1. The highest increase in activity is shown for a concentration of 14 %. Our theoretical investigations are not modeling the concentration ranges explicitly but rather focus on the underlying mechanism of improving activity.

4.5.2. Role of coverage and the inactive cus site

The computational results show promise but other computations have shown not so much promise as indicated here. As an example the results of a computation for the OER reaction steps [Eqs 4.1.1-4.1.4] are now compared to the previous computation. The difference is only in the occupation of the nearest cus site. Our computations have assumed an oxygen atom at this site because of the phase diagram analysis that was performed in section 3. Historically many computations are performed with an empty cus site so we add it for comparison now. This simple example from figure 4.5.4 shows the importance of careful considerations when setting up computational analysis models and also shows the sensitivity of the OER to surface configurations.

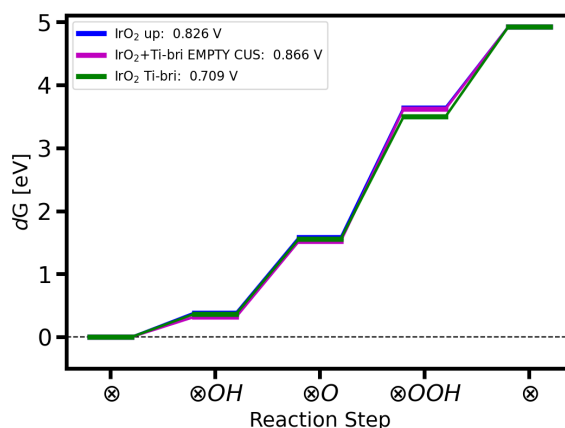


Figure 4.5.4.: Gibbs free energy reaction steps on Ti-mixed IrO₂ in bridge position with and without an oxygen atom in the nearest neighbor site. The stabilizing effect has vanished in the scenario of the empty cus site.

4.6. Results

From the experimental results we know that mixing IrO₂ with a low amount of titanium improves activity. This result is not directly explained when comparing to the computational results but when examining the reaction path closely we can see that there is a discrepancy when computing the pure iridium overpotential. The reaction equations do not describe the correct adsorption path on the IrO₂ (110) rutile surface but ignore the deprotonation effect of the *OOH adsorbate. When taking the correct adsorbate configuration into account we reach the same conclusion for the bridge-stabilized Ir-Ti motif in figure 4.5.3. This configuration improves the overpotential by about 120 meV which is in agreement with the experimental results. The correct description of the adsorbate configuration is critical to capture this effect [114]. Furthermore, a more in-depth analysis of possible changes of the adsorbed species on the neighboring CUS site needs to be investigated in more detail. Possible processes involving the adsorbate on the neighboring CUS site may have a more influential effect than has been indicated so far. The route of replacing the host species might present a general way forward in designing new catalysts by choosing a base material which has optimal adsorption energy steps close to 1.23 V per step and offers the ability to tune the *OOH step towards the optimal energy.

5. Hydrogen coverage of the IrO₂ Surface

To obtain statistical data for large systems we can not apply standard DFT. Computing a periodic structure containing more than 2000 atoms can be achieved today but the computational cost is high. If it was only a single structure to be computed this would perhaps be reasonable. But for the assembly of statistical data we need many structures and by that the direct computation by DFT is not feasible. The central quantity in thermodynamics is the partition function

$$Z = \sum_i^{\infty} \exp \left(- \frac{\sum_m (E_i - \mu_{i,m} N_{i,m})}{k_B T} \right), \quad (5.0.1)$$

where $\mu_{i,m}$ and $N_{i,m}$ represent the chemical potential and the number of constituents of species m from structure i , respectively. k_B denotes the Boltzmann constant and T the temperature. While this formula is analytically exact and one can compute energy values for $E_{i,m}$ and $\mu_{i,m}$ from DFT there is no numerical application for this formula to the current problem. In this form there is no applicable criteria one could use to determine a convergence of this infinite sum. Educated guesses from physical or chemical intuition might help constructing a partial sum but there is no mathematical way of proving that the constructed sum represents a real structure. Applying methods of statistical mechanics to problems with many degrees of freedom requires a different theoretical approach. A solution is the cluster expansion as presented in chapter 2.

5.1. Defining the Problem

For a cluster expansion we need to define a primitive input cell. The cell is the basis for the cluster expansion. It contains all structural information about the problem to solve. In it we define the atomic positions and species of the structure as well as possible vacancies or substitute species. For the particular problem we investigate there are three possible positions in which hydrogen can be replaced with a vacancy on the surface. All other species are constant. We are using the surface we have previously defined as 1x1 here in contrast to the 1x2 we have employed in the previous chapters because the 1x2 is not a primitive cell but a simple repetition of the 1x1. For a consistent cluster expansion one should use the smallest possible representation of the basic input cell, i.e. the primitive cell. The possible hydrogen / vacancy positions are shown in figure 5.1.1. Possible positions for hydrogen and vacancies are limited by the physical space on the surface. When adding too many hydrogen atoms they will not adsorb onto the surface but form H₂ molecules that desorb into the vacuum resulting in a structure that is not representative of the defined problem. Testing for possible errors in the starting structure that result in unusable structures needs to be done carefully. The CASM code then creates all possible symmetrically distinct structures from this primitive input cell up to a specified cell volume. The smallest set of structures contains all possible surface configurations of the primitive input cell geometry. For our model representation the surface structures are relaxed with the constraint that the lower half of the structure is fixed to bulk positions while the upper half with possible hydrogen can fully relax. The criterion for a converged structure is set to reach all residual forces below 0.01 eV/Å. The first 97 structures of all 282 structures were created by a different cluster expansion code. At the time of the start of the project the CASM code was in its infancy and we started using a different code, ATAT [115, 116]. While the theory is identical the technical implementation of it is very different be-

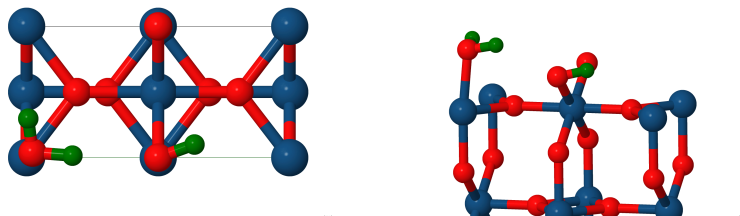


Figure 5.1.1.: Primitive input cell where green positions are possible vacancies or hydrogen positions that will be populated automatically during the input structure generation by the CASM code.

tween CASM and ATAT. CASM creates directly enumerated structures of which the user can select a subset to compute. ATAT generates structures by a different algorithm so the resulting structures do not have to be directly enumerated. For this reason there are some structures of lower cell volume not included while the cluster expansion includes structures up to volume 5. A predictive method of which structures give the most information to the machine learning process does not exist as of this writing. There is work done to improve on prediction of structural information but this is difficult. Predicting the information content of an unknown structure is only possible by learning from what information we gained from the previous input structures. For some problems these algorithms may work while they fail for other structures so that one is left with the only reliable method, direct enumeration. Direct enumeration is very successful for problems with low degrees of freedom in the species but can prove computationally inaccessible for other problems. In a case of 16 possible hydrogen / vacancy positions one ends up with 16384 structures to compute for the primitive input cell alone when using direct enumeration. Other strategies can be used to tackle these problems within a cluster expansion framework.

5.2. Generating Structures for Training Data

The structures we compute form the basis for our cluster expansion. From them we use the total energies and the atomic positions are informational input. The CASM code takes some simple commands to create a cluster expansion. Once the primitive input cell is defined the user has to issue a few commands which are shown in the appendix II for the interested reader.

5.3. DFT code details

All input data for the cluster expansion was computed with FHI-aims. We use a two step relaxation procedure with first light and then tight basis settings [69] with the RPBE [46] functional. The **k**-grid was set such that the projected grid of the primitive cell is at least equivalent to $10 \times 10 \times 1$.

5.3.1. Matching relaxed structures

During the import of structures for the cluster expansion they are checked for symmetry. Only one symmetrically distinct structure representation exists for each configuration in the cluster expansion. For different reasons there can be multiple structures that have distinct starting configuration but after relaxation show the same final configuration. Here we show three different reasons from three examples of this procedure. Reasons are for the first one technical and physical, the second one configurational and the third one physical only. In all examples the starting structures are symmetrically distinct. The default procedure in CASM when duplicate structures are detected is to resolve the problem by accepting the structure with the lowest energy in this configuration.

Mirror Symmetry

The orientation of the hydrogen is the difference in starting configurations in this example. From previous computations we already know that the energetically preferred orientation for the hydrogen atoms is the cus direction. For the cluster expansion however, this knowledge can not be exploited during the generation of starting configurations. During relaxation the hydrogens change their position from the less favorable in-bridge orientation towards the known in-cus orientation. After relaxation the relaxed structures have mirror symmetry which is detected by CASM. The difference in configuration is only the orientation of the hydrogen at the bottom of the picture, to the left and to the right, as shown in figure 5.3.1. However, when the structure is mirrored in the middle we see the symmetry. This results in a duplicate structure for the same configuration. For the cluster expansion the configuration with the lowest total energy is taken as representative and the other results are discarded.

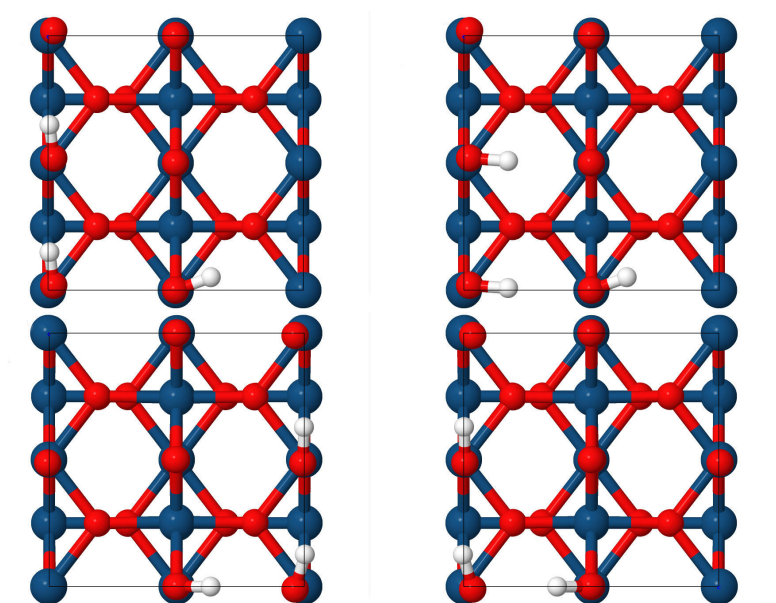


Figure 5.3.1.: Mirror symmetry example: Upper left and right: Unrelaxed DFT input, bottom left and right: final DFT relaxed structures

Positional deviation

As we have seen in the previous example there is a large change in configuration accompanied with symmetrical changes. The water molecule visible in the upper left picture in figure 5.3.2 dissociates. One hydrogen moves towards the bridge position. The relaxed structures are detected to be equivalent because of a positional deviation from the primary input cell definition. Remembering figure 5.1.1 we see there are three possible positions for vacancies / hydrogens. Examining figure 5.3.2 bottom left, we see the hydrogens to the lower left and lower right in the picture have relaxed into positions that are defined in the primary input. They represent the in-cus directional position of the water hydrogen from the primary input. In contrast to this we see in the bottom right of figure 5.3.2 also two hydrogens but these have relaxed into a position that is not defined in the primary input. This is handled in CASM by finding the nearest possible position that resembles the relaxed configuration most accurately which in this case translates into the same final relaxed configuration. This misrepresentation of a defined position in the primary input is resolved also by including the structure with the lowest total energy while discarding the other. This behavior can become a problem for a cluster expansion in two ways:

- The final positions of the relaxed configurations are too far from defined positions:
The cluster expansion can not identify the configurations correctly and does not resemble the actual structure description. Thus the cluster expansion describes a different structure than is actually defined in the primary cell.
- Changing the primary input so that it includes many possible positions:
This will result in many degrees of freedom for the cluster expansion. While this will help reducing the previously described

issue of positional uncertainty it equally introduces computational complexity. This results in problems defining a reasonable basis set and fitting parameters during later steps and will negatively impact computational performance for the whole cluster expansion.

If these problems arise during a cluster expansion a solution is in most cases found by trial and error. There is no analytical convergence criterion to identify the ideal number of vacancies / positions one needs to define. While this problem is for most bulk cluster expansions not relevant it can become a burden for a surface cluster expansion.

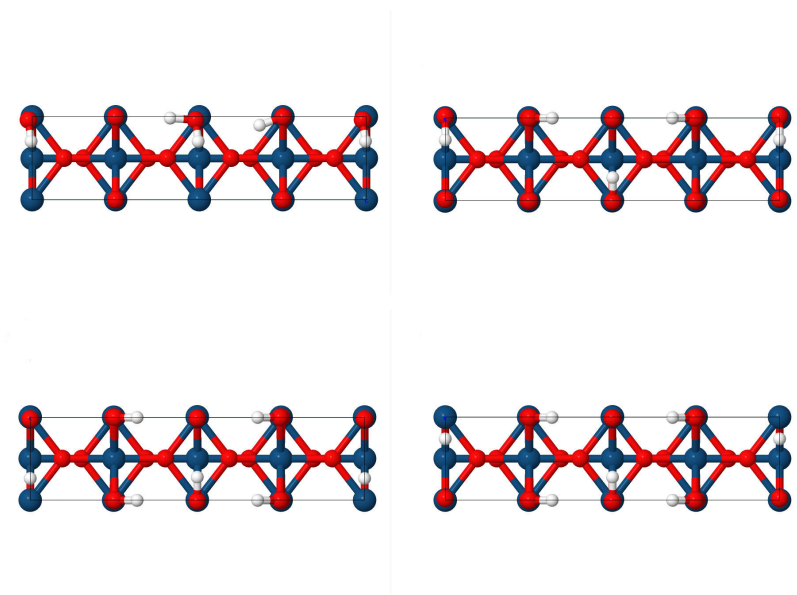


Figure 5.3.2.: Positional deviation example: Upper left and right: Unrelaxed DFT input, bottom left and right: final DFT relaxed structures

Favorable Positions

The relaxation in this example has purely physical reasons. Both defined unrelaxed configurations in figure 5.3.3 top left and right are distinct by the hydrogen positions in the center of the picture. During relaxation the in-bridge oriented hydrogen relaxes to the in-cus direction as this configuration is energetically more stable. The result is an identical configuration which is detected as such. The lowest total energy is taken and the other is discarded.

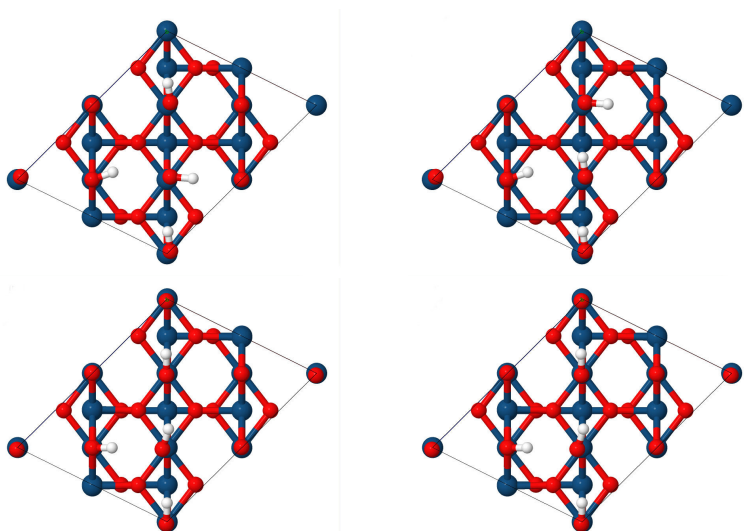


Figure 5.3.3.: Favorable position example: Upper left and right: Unrelaxed DFT input, bottom left and right: final DFT relaxed structures

5.4. Grand Canonical Phase Diagrams

Within the same setting as in the previous chapter we can plot the surface phases for each structure of the cluster expansion input. The result in figure 5.4.1 is confusing when one tries to find out single stable phases. The plot is not meant to be analyzed as before but is rather presented to show that there are many competing phases present on the surface and not simply one single stable phase. Also we can see that some configurations of hydrogen coverage are not favorable. The reference is chosen the same as in the previous chapter. Parallel lines

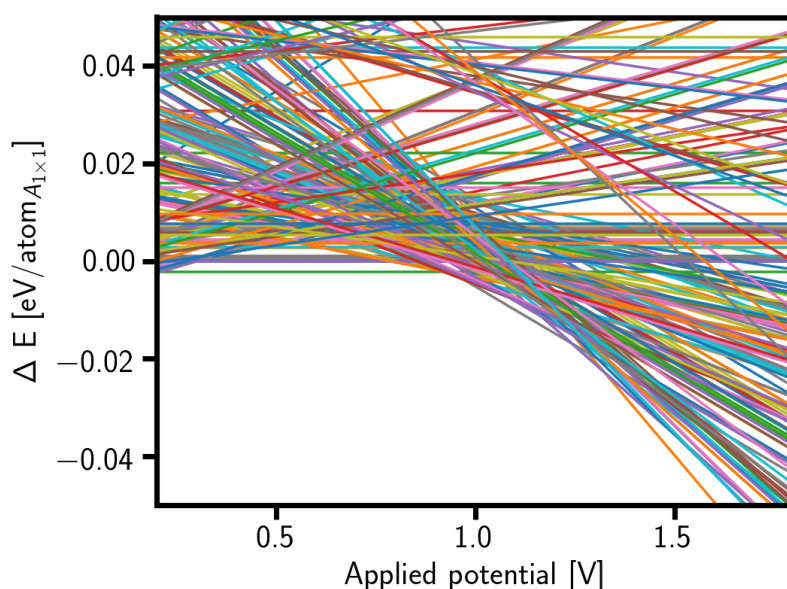


Figure 5.4.1.: Surface phases of all cluster expansion input structures constructed by the same rules as in the previous chapter. This is meant to illustrate the lot of competing phases with same composition (parallel lines) that are included in the cluster expansion.

indicate same compositions of different configuration. When looking for parallel lines to the reference (zero-line) we can see that we did find new structures of lower energy than we did find before in our 1×2 simple surface model from the previous chapter. The previously assumed ground state and the newly found ground state for the water covered surface are shown in figure 5.4.2. While the configuration of the 1×2 ground state is also included in the cluster expansion we see that the more stable configuration is on a 2×1 multiplication of the primitive cell. This configuration is formally compatible to compute a phase diagram but would have needed to be guesses intuitively during the work in the previous chapter. In direct enumeration the cluster expansion covers possible structures in a systematic way and we effortlessly find more favorable configurations. Intuitive guesses might be helpful in some cases but are certainly no match for a directly enumerated search as shown here.

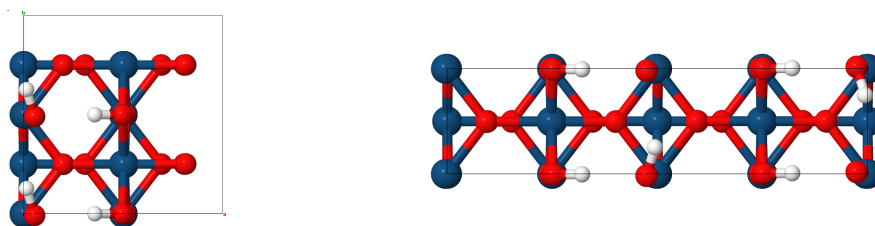


Figure 5.4.2.: Ground states of the water covered surface from 1×2 phase diagram analysis (**left**) and cluster expansion (**right**)

To compare the surface phase diagrams further we now separate out the phases of same composition and plot only those of lowest energy. This way we can compare the results directly and see whether we actually did find new phases. The plot analysis is shown in figure 5.4.3. In this plot the phases from the previous chapter are shown as dotted lines and labeled with prefix GCPD (Grand-Canonical Phase Diagram). The configurations from the cluster expansion in figure 5.4.3 are labeled by numbers indicating the number of the structure within the set of the cluster expansion input data. There is no physical meaning behind this number, it is simply direct enumeration meant as unique identification. We use the same reference as we used before in the 1×2 surface phase diagram in the previous chapter. We can immediately see that we did find a configuration with lower energy in the cluster expansion as already shown in figure 5.4.2. The phase of structure 22 is parallel to GCPD $\text{E}(*\text{H}_2\text{O}*\text{H}_2\text{O})$ but lower in energy indicating a more stable phase. The next phase represented by structure 54 and GCPD $\text{E}(*\text{H}_2\text{O}*\text{OH})$ is energetically very close. The phase 185 is not present in the static 1×2 phase diagram because the composition can not be represented on this surface. Phase 27 and GCPD $\text{E}(*\text{H}_2\text{O}*\text{O})$ are of same composition again but phase 27 is energetically much lower than GCPD $\text{E}(*\text{H}_2\text{O}*\text{O})$. Phase 29 and 1 and then again crossing at about 1.3 V. Although we have a lot more structures included in the cluster expansion, some of which form phases more stable than we have previously seen, the crossings of the phases are very similar to the results we have seen before. The phase crossings of the first three structures 22, 54 and 185 are in the middle of the wide first peak of the experimental CV and the phase crossings of 27, 29 and 1 are at about 1.3V, similar to the previous analysis. From this analysis we did not gain a different insight as to what the factors behind the shapes of the peaks are, seen in figure 5.4.3. To gain further insight we utilize the cluster expansion now.

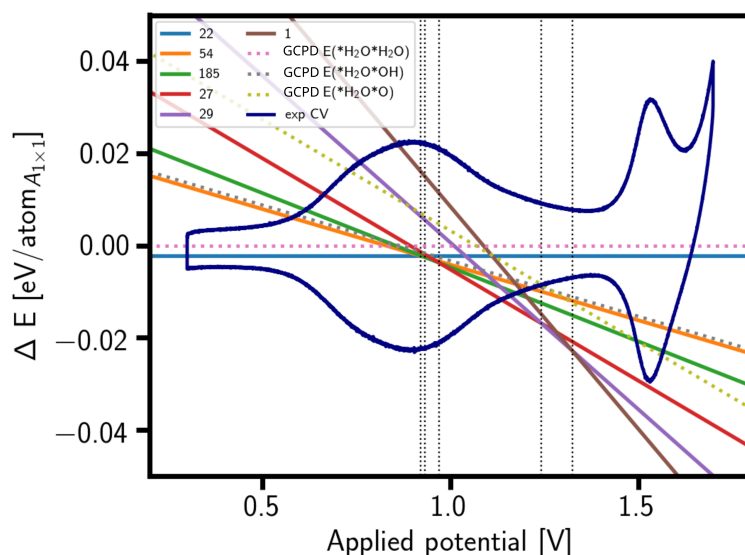


Figure 5.4.3.: Comparison of previously found DFT ground states on the 1×2 surface (dotted lines, labeled GCPD) to lowest structures from the cluster expansion input and the experimental shape of the cyclic voltammetry. Transition energies roughly coincide with experimental data but the peak shape of the experiment cannot be explained.

5.5. Hydrogen chemical potential alignment

The CASM code sets the chemical potentials of the included species automatically. When expanding bulk materials this procedure is reliable and usually not error prone. In this project we investigate the hydrogen coverage on the (110)-surface of rutile IrO₂. When hydrogen is stripped off the surface it forms H₂ gas under electron exchange at

the cathode. Thus the gaseous hydrogen is our reference. When using the automatic setting from CASM this is not accounted for and we need to apply the relevant correction to the internal CASM chemical potential to ensure compatibility to our externally applied potential. The internal chemical reference in CASM is computed from

$$\mu_{\text{H}}^{\text{CE}} = \frac{1}{N} (E(N \cdot \text{H}) - E(0 \cdot \text{H})), \quad (5.5.1)$$

where $N = 3$ is representing the number of hydrogens that can be adsorbed at maximum in the primitive unit cell we defined as basis for the cluster expansion. The hydrogen gas reference we compute externally in computed from

$$\mu_{\text{H}}^{\text{ref}} = \frac{1}{2} (E_{\text{H}_2}^{\text{DFT}} + \text{ZPE} - TS_{\text{H}_2}) - U_{\text{RHE}}, \quad (5.5.2)$$

which is the same as we used before in equation 3.6.2 during the construction of the surface phase diagrams. It contains corrections for zero points vibrations and entropy contributions. The calculated difference of 0.368 eV between $\mu_{\text{H}}^{\text{CE}}$ and $\mu_{\text{H}}^{\text{ref}}$ is the shift we need to apply to correct the internal CASM reference to our external potential reference U_{RHE} .

5.6. DFT Formation Energies

The input to our cluster expansion are the relaxed structures and their respective total energies. From this data CASM computes the formation energy per unit cell of each structure. We can display all our input data sorted by composition to gain an overview of the distribution of data. This is shown in figure 5.6.1. The x-axis represents the coverage of hydrogen on the surface while the y-axis shows the formation energy per unit cell in eV. This data has been prepared for the fitting process already in the sense that we only include configurations that have formation energies lower or equal to zero. In other

words we only include stable structures. DFT ground states are plotted as solid dots and connected by a line forming the convex hull of our input structures. Other formation energies are plotted as open circles. From figure 5.6.1 we can see a few characteristics already. First, the most stable composition has a hydrogen coverage of 2 resulting in a water covered surface. We have seen this before from the phase diagrams. At the same composition there are many different possible configurations with a wide spread of energy. This means that the hydrogen on the surface prefer some configurations above others. When we think about the primitive input cell again (c.f. 5.1.1) we have three possible hydrogen / vacancy positions. Structures that contain three hydrogen per primitive cell are energetically unfavorable while structures with two hydrogen per primitive cell are close to the ground state. Forming super cells with islands of 3H and 0H coverage per primitive cell is possible by construction in this composition range but not favored by energetics. The same principle applies to all structures with lower hydrogen coverage than two. Further we see that in the input data we computed there is not for every composition a DFT ground state. Computing more input data might reveal ground states but this is not a necessity. These states may be transition states between ground states. Computing more input data could help clarify this observation. To make use of the grand canonical Monte-Carlo simulation from CASM we need to find a representation of our DFT input data set. This representation will be used to predict formation energies of larger unit cells than we have used to compute the DFT input data. To achieve this we make use of machine learning methods we apply to our data set. We train a model that is used to predict the formation energies of different compositions and configurations.

Cluster Expansion Basis Settings

As we have seen in the theory chapter there are parameters we need to set in order to fit a predictive model to our data. Explicitly, the

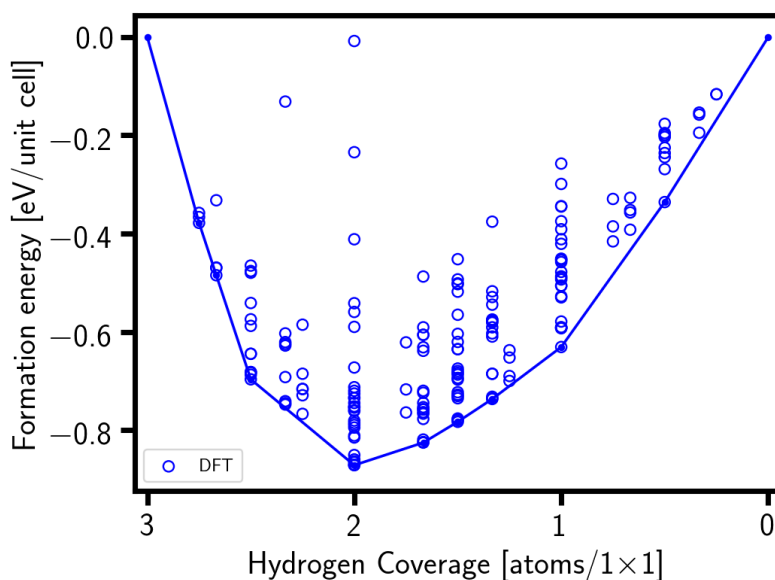


Figure 5.6.1.: DFT formation energies of the input structures to the cluster expansion and their convex hull. Regarding the energies at coverage 2 we can see that not all water structures are indeed favorable but some show unfavorable configurations in terms of energy.

employed basis set is defined by

```
"basis_functions" : {
  "site_basis_functions" : "occupation"
},
"orbit_branch_specs" : {
  "2" : {"max_length" : 13.0},
  "3" : {"max_length" : 9.0},
  "4" : {"max_length" : 7.0}
}
```

This defines the basis used for fitting the DFT data with pair interactions up to range 13 Å, triplet interactions up to 9 Å and quadruplet interactions up to 7 Å. We select `occupation` here because we can easily visualize the clusters and their respective ECIs. The `chebychev` variant does not change results in our case. In simple words we want each open circle in figure 5.6.1 to be filled with a red dot and the convex hull to be represented as well as possible. For the mentioned parameters we obtain figure 5.6.2. In this we can identify a few things. There are some deviations at coverages above 2.5 hydrogen per unit cell, towards the left side of the graph. For the current project we regard these as irrelevant because we are not interested in this coverage range. The range from two towards zero coverage is our focus here. There are also small deviations in ground state representation in this range. One ground state at coverage 1.5 is not represented on the convex hull but the model prediction is very close while the prediction at 1.25 coverage is a bit too stable compared to the DFT data. A small prediction error is also visible for ground states at coverage 1 and 0.5 while there is a predicted ground state at 0.3 coverage. Overall the present states in the DFT data are well represented by the model. Deviations of 100 meV in formation energy per unit cell translate to an accuracy of $100/34 \approx 3$ meV per atom since we have 34 atoms in the primitive input cell.

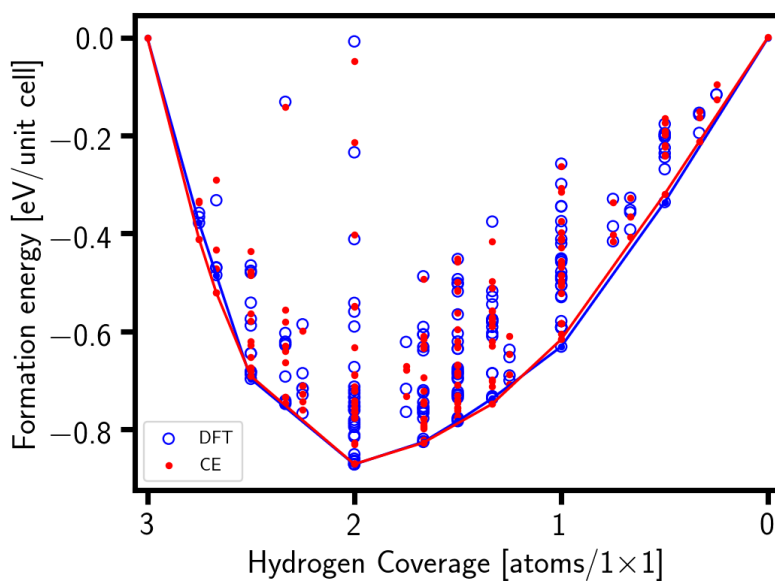


Figure 5.6.2.: DFT formation energies and fitted representation in the cluster expansion. The convex hull is well represented with a little misfit to the left side. We disregard this error because we are not interested in this composition region. Also not every single open circle is represented by a red dot but the errors in energy are acceptable to use this fit as a representation for our system.

5.7. Clusters in the Expansion

Effective cluster interactions (ECI) are computed from the DFT input configurations. The maximum number of clusters thereby depends on the input configurations. Larger super cells yield more diverse composition ranges and with that more possible clusters. Within CASM the ECIs of the clusters are computed at the start of the fitting procedure. The values determine favorable or unfavorable configurations of clusters of which the predictions of total energy are later based. The number of actual clusters included in the cluster expansion is determined by the basis settings. Depending on the choice of basis set CASM includes only the cluster specified in the settings. In the current cluster expansion there are pairs, triplets and quadruplets included. Two examples of each group are randomly shown here to represent their family. In total there are 29 pairs, 59 triplets and 14 quadruplets included in the current expansion with the present basis settings. These numbers are symmetry dependent.

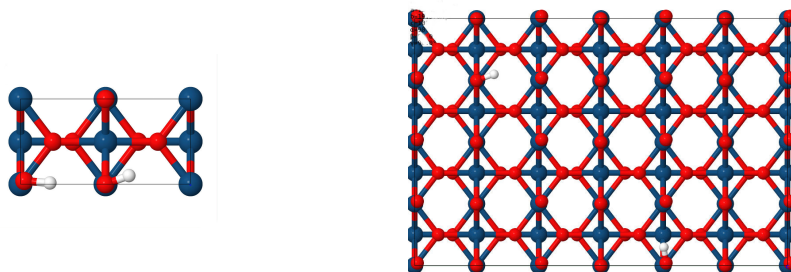


Figure 5.7.1.: Example configurations for pair interactions, cluster 1 and 27, ECI 0.519 eV and -0.693 eV, respectively

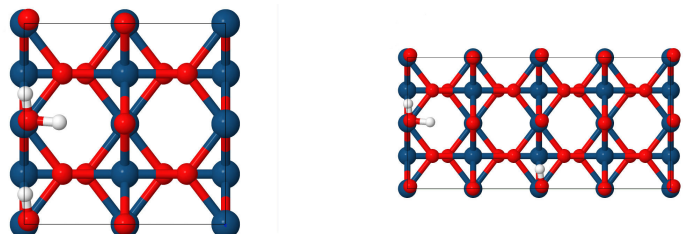


Figure 5.7.2.: Example configurations for triplet interactions, cluster 1 and 26, ECI -1.549 eV and 0.239 eV, respectively

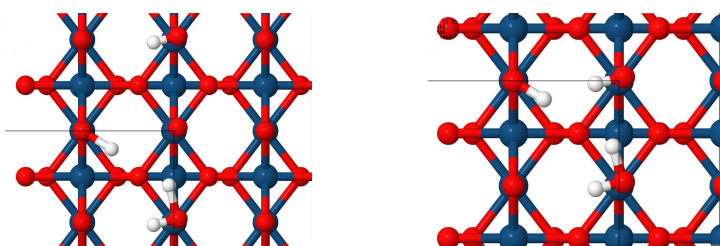


Figure 5.7.3.: Examples configurations for quadruplet interactions, cluster 2 and 7, ECI -0.358 eV and -1.475 eV, respectively

5.8. Monte-Carlo Simulations

Since DFT is a zero-Kelvin theory the convex hull applies strictly only when the temperature is at 0 K. At higher temperatures there is a finite probability that states of same composition but of other configuration than the ground state will be realized. This probability depends on temperature and energy difference to the convex hull ground state. The algorithm which determines the probability is the established Metropolis algorithm [90] which is commonly applied for Monte-Carlo simulations. With the model we have trained we can

now predict formation energies of structures that we can not compute within standard DFT we have used to compute our training data. The input file for the Monte-Carlo simulation and the parameters are shown in detail in appendix III. To ensure converged results from the Monte-Carlo simulation with respect to the simulation cell size the simulation was run with different cell sizes. We picked sizes of 40, 60, 80, 100 and 120 to be interpreted as the primitive unit cell multiplication factors in the (a,b)-plane. A size 40 run stands for a simulation with a cell size computed by

$$\begin{pmatrix} 40 & 0 & 0 \\ 0 & 40 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix}, \quad (5.8.1)$$

which results in a cell dimension of $128.377 \times 258.377 \text{ \AA}^2 = 12.838 \times 25.838 \text{ nm}^2 = 331.696 \text{ nm}^2$ for the smallest simulation. Cells of this size are not accessible by standard DFT today. The largest cell of 120 results in a surface area of 2985.263 nm^2 . The resulting plot is shown in figure 5.8.1. We can see that the Monte-Carlo simulation of the smallest cell already yields quantitatively the same results as the simulation from cell size 120. The only difference is the noise which reduces strongly with increasing cell size. We use the results from the 120 cell regarding results from Monte-Carlo simulations.

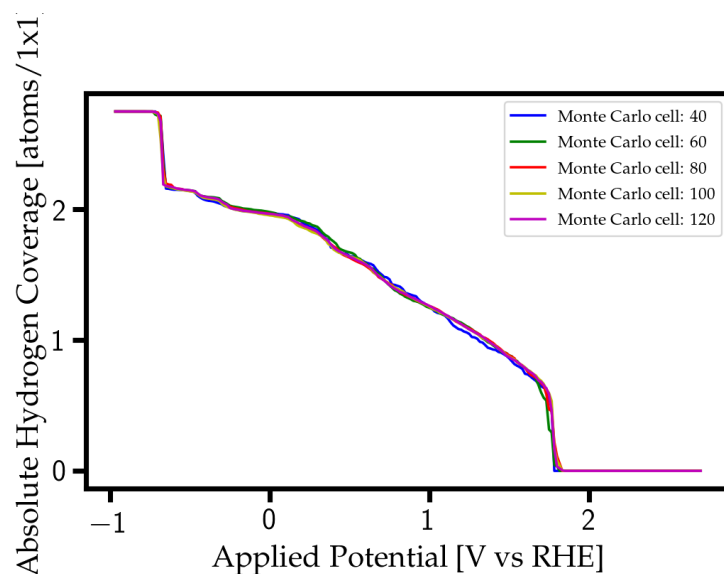


Figure 5.8.1.: Monte Carlo Convergence Test. Results are similar when comparing different simulation cell sizes and the noise reduction is clearly visible. In this plot the complete range of simulated potential is shown from -1 up to 2.8. This is to demonstrate that the simulations all have the same starting and ending conditions that do not change with the simulation cell settings. Later we will show only the potential range that we are accessing experimentally, the range from zero potential up to 1.8 V.

5.9. Adding Zero-Point Energies

Up to this point the DFT input data and with that trained model used for prediction is based on pure DFT total energies. As we have seen in the grand canonical phase diagram chapter, zero point energies have an influence regarding the energies of adsorbate species. To include the effects of zero point energies in our cluster expansion model we follow a simple procedure. For each DFT input structure the adsorbates are classified into the groups as we have done before. In this case we only have OHHc, OcHb, OHc and Oc adsorbates. With a simple script the adsorbates are counted and the zero point energies of each respective species are added to the total DFT energy. After that the cluster expansion is performed again with the same structures but with modified energies that now contain the zero point energy contributions. As before we set the entropy of adsorbed species to zero. We repeat the fitting process with the parameters mentioned in appendix IV and the result is shown in figure 5.9.1. There are only minor differences comparing the DFT+ZPE input data to the DFT-only result from figure 5.6.2. The training data set is well represented again and we can run Monte-Carlo simulations with the model trained with the data. The result is shown in comparison to the non-ZPE Monte-Carlo results and the experimentally derived coverage in figure 5.10.3. From this result we can quickly see some differences we will analyze further in the next section. Major differences are:

- The coverage at zero potential is lower due to the increase in energy from the zero point energies. This is physically meaningful since at finite temperature we have more mixing between the DFT ground state and configurations of similar composition but higher energy resulting in compositions being represented that have lower hydrogen coverage.
- The strip-off potential has changed to a lower value. This is also meaningful because with higher vibrational energies the

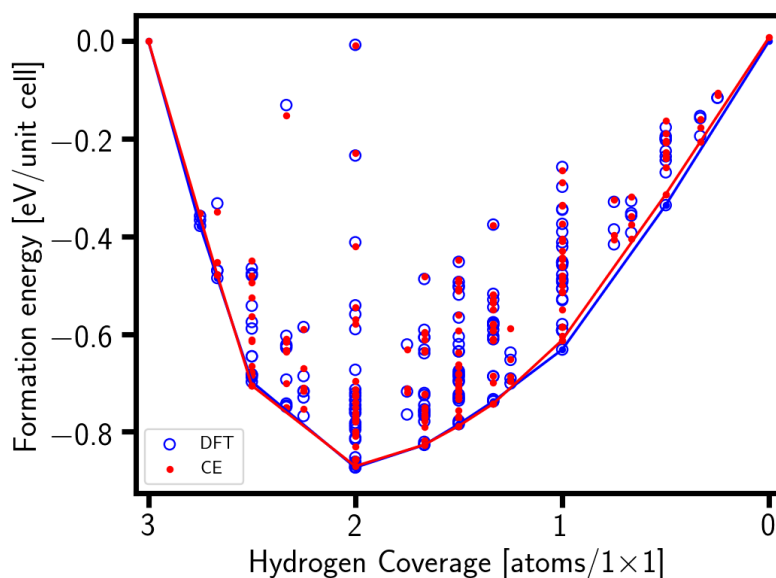


Figure 5.9.1.: Results from the fitting with ZPE added energies. Minor differences are visible on the left and right flanks of the convex hull as well as different representations of individual structures. Overall we achieve a suitable representation again and use it for Monte-Carlo simulations.

hydrogen can desorb easier from the surface.

- The number of hydrogen atoms that desorb up to the steep drop in coverage has changed. While the result without zero-point energy correction suggests about 1.2 hydrogens desorb the zero-point energy corrected Monte-Carlo simulation suggests 0.5 hydrogens desorb.

Overall the addition of zero-point energy correction improves the description of the hydrogen coverage with respect to the physical and

chemical properties that are defining the process.

5.10. Interpreting the Monte-Carlo Results

Comparing the results of the Monte-Carlo simulations without added ZPE and with ZPE [figure 5.10.1 top vs figure 5.10.1 bottom] we can see major differences with respect to the species present at the same external potential. The species present at zero external potential are different in quantity. For the non-ZPE Monte-Carlo results the surfaces with keys clean, (1), (3) and (5) are present with only less than 5% of total coverage. The fully covered surface is at about 15% while (2) and (4) are at about 18%. Structure (6), the water covered surface, is dominating with 50%. The phase separation is showing clearly at 1.7 V vs RHE. These results agree very well with the previously performed surface analysis from the phase diagrams from which we saw a strong stability of the purely water covered surface. Due to the added entropy by the Monte-Carlo simulation the part of the surface which is not covered by the structure (6) (water covered) is almost evenly shared between the fully covered, the deprotonated water (4) and a single hydrogen structure (2). We see that non-ZPE results clearly favor single stable phases while others of the same composition as (2), i.e. structure (3) are almost absent. With regard to the manifold of energetically similar input structures we see in figure 5.4.1 this result is physically not strongly convincing. Looking at the results from the Monte-Carlo simulation with added ZPE we see that the coverage at zero potential has changed. There are now more different types of configurations present on the surface. Structure (6) has lost presence from almost 50% down to about 5% and with that the fraction of the surface that is covered by water molecules is now very low. The partial coverage of structure (4) remains almost constant at about 18% while other configurations that were not present before all cover a

part of the surface now. The most significant change occurred in the clean surface. While not present during the non-ZPE Monte-Carlo simulation it has shifted towards 20% of the surface coverage. The fully covered configuration of the surface has also risen and makes up most of the individual configurations now. The phase separation has shifted towards lower voltages. In total we can see more noise within the Monte-Carlo results due to the added amount of entropy in the input structures. From a physical standpoint this is reasonable because more entropy results in more free energy which then results in lower energy differences between configurations and with that a broader mix of configurations.

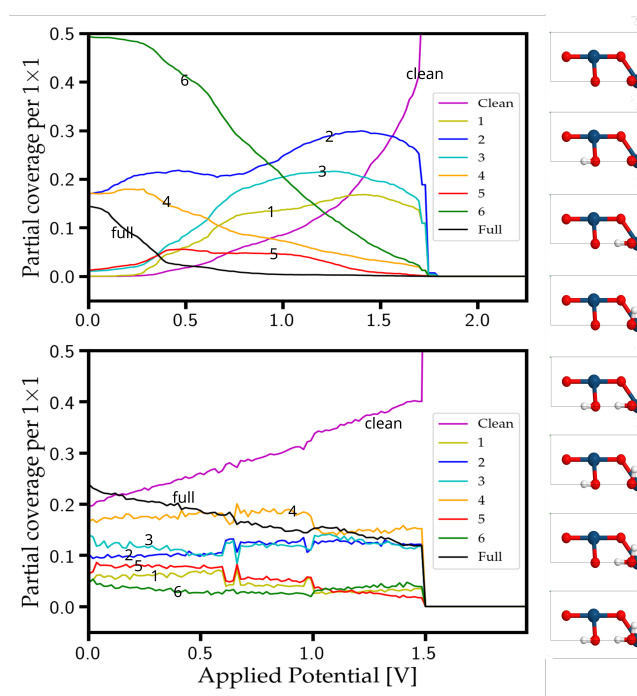


Figure 5.10.1.: Configuration resolved coverage depending on $U(\text{RHE})$ and numbered structures of species. **top:** Coverage for the non-ZPE results, **bottom:** Coverage for ZPE added results. **Right:** Occurring surface configurations in the same order as listed in the plot labels from top to bottom. The coverage plots are cut at 0.5 maximal coverage for better visibility, above 1.75 V in the top and above 1.5V in the bottom plot the only present configuration is the fully covered surface (magenta line reaches 1).

5.10.1. Computational CV

From the experimental data we have a CV for our surface. This CV is the first derivative of the coverage. Now that we have computed the coverage of hydrogen on the surface we can from that get the CV. The results are shown in figure 5.10.2 for the zero-point corrected Monte-Carlo results. A conversion was performed to scale the computational results to the experimentally measured units of electrons per surface area. For this purpose we count the number of electrons transferred, which is equivalent to the number of hydrogen atoms desorbed, and divide by the surface area of our Monte-Carlo simulation cell. Then we can compare the result directly to the experimentally measured CV. In this plot we see that we are missing the physical description responsible for the first wide peak in the experimental CV. The second peak is well covered. The underlying theory that generated the input data for the cluster expansion is not able to describe the first peak. Possible reasons are:

- We do not have an explicit water layer above the adsorbed hydrogen on the surface. Water molecules are highly mobile and can contribute to more favorable surface energies by donating partial charges towards unsaturated surface charges or hydrogen atoms.
- An implicit solvent model might be able to partially produce results from an explicit water layer. But explicit water layers on the surface followed by an implicit solvent background might be the physically most accurate description possible today [117]. The computational cost of a cell with water layers is hard to estimate because of the many degrees of freedom that water presents in such an environment. A full cluster expansion with the number of training data structures as used in this project might not be achievable in reasonable time.
- Vibrational interactions between adsorbed hydrogen on the sur-

face are neglected. The vibrational energy of water might change depending on the neighboring adsorbate, these contributions are not taken into account here. Zero-point corrections are applied only on the basis of local types of adsorbates and terms of higher order are neglected.

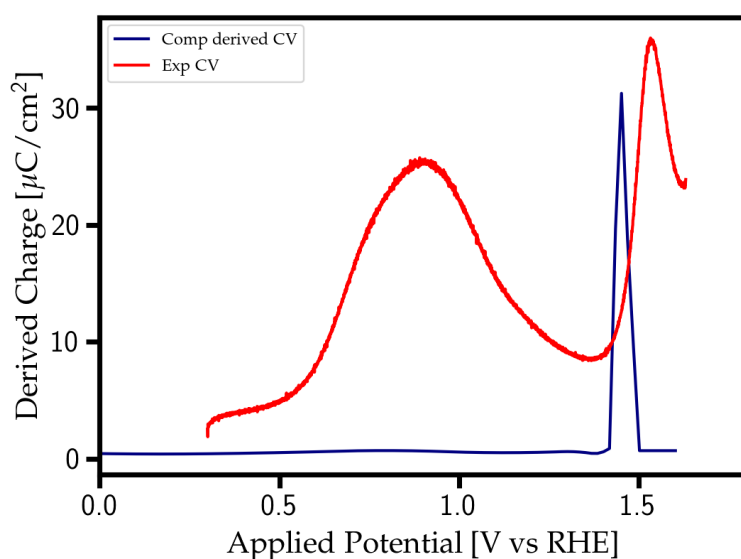


Figure 5.10.2.: Computational CV from the Monte-Carlo simulated coverage compared to the experimental data. The first peak is not described while the second peak matches reasonably well.

5.10.2. Monte-Carlo snapshot analysis

As the applied potential gradually increases the hydrogen coverage decreases linearly. From the snapshots of the Monte-Carlo simula-

tion there is no preferred hydrogen configuration on the surface during this process. The mixing of states on the surface due to entropy is dominating the region up to 1.5 V for uncorrected and 1.3 V for zero-point corrected computations. After this point there is a steep drop in hydrogen coverage which hints to a change in the underlying mechanism. When looking at snapshots from the Monte-Carlo simulation in this region we can clearly see islands of hydrogen appearing which minimize the region where OH adsorbates and the clean surface are adjacent. Detailed snapshots are in the appendix, section VI.

5.10.3. Comparison non-ZPE vs ZPE vs experiment

Finally a comparison of the Monte-Carlo results with and without ZPE can be made to experiment in figure 5.10.3. In the lower left corner we see the experimental CV with the number of electrons exchanged during the first process captured by the first peak. The integrated coverage from the experimental CV is shown for the measured voltages in the plot, too. A full measurement over the whole simulation range from 0 to 2V RHE is not available. The number of electrons is integrated to 0.85 [15]. Since our simulation can not capture the first peak we assume the process to stretch continuously from 0V vs RHE up to the phase separation point at 1.7 V and 1.5 V, respectively. The total exchange of electrons during the experimental first peak phase is equivalent to the number of stripped hydrogen from the surface of our simulation. This count results in ~ 1.2 electrons from the non-ZPE simulation and in ~ 0.5 electrons in the ZPE Monte-Carlo simulation.

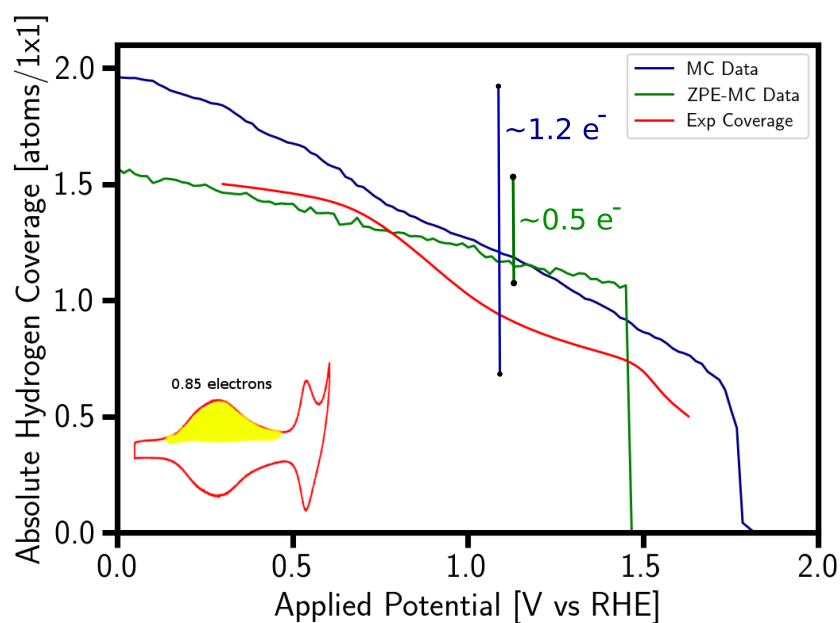


Figure 5.10.3.: Over- and underestimation in exchanged electrons compared to experimentally integrated number. We see that non-ZPE results overestimate the number of electrons that are exchanged while results with ZPE suggest a smaller number than from experiment.

5.11. Results

From the results presented in this section it becomes clear that we are missing at least one physical effect to describe the experimental cyclic voltammetry. From the applied standard DFT we can not capture the first peak. The underlying effect might be related to solvation interaction effects that are not modeled in this study [118]. With recent developments in DFT codes like jDFTx and the solvation models implemented in this code it is in principle possible to investigate these effects from a theoretical point of view. In practice however, one needs to perform a complete cluster expansion again which for the presented problem of the hydrogen covered IrO₂ surface would include roughly 300 DFT computations. Also we can not model pH effects as these are not included in the current level of theory. Furthermore, effects of ions at or near the surface are also neglected. These ions stem from the adjustment of pH levels during experimental measurements with basic solutions as NaOH, KOH or LiOH or acidic solutions like H₂SO₄ [15]. The second peak is well described with our applied method. The physical effect of forming anti-bonding atomic orbitals in the O-OH configuration is captured by standard DFT. Compared to a similar computational analysis by Karlberg [119] we find more complex physical properties on the investigated IrO₂ surface than are at play on a pure platinum surface.

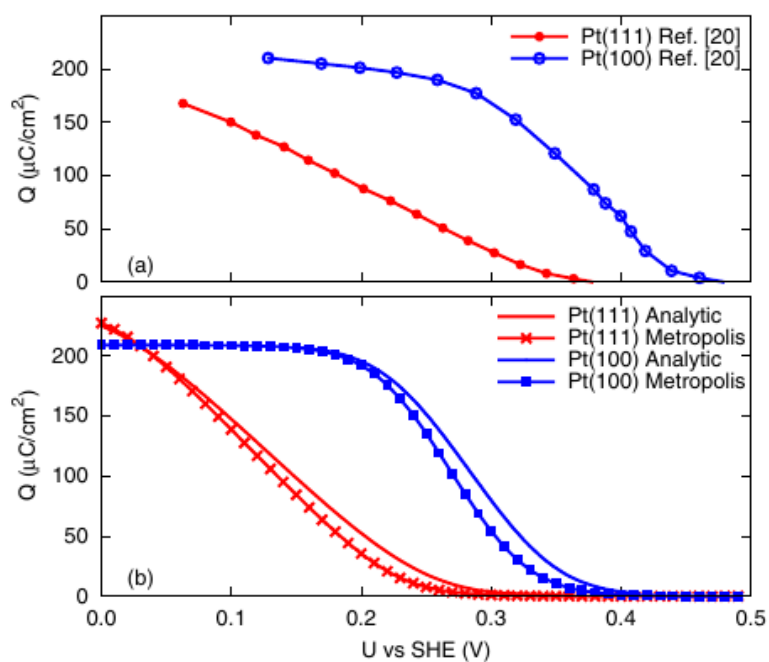


Figure 5.11.1.: Results from Karlberg cluster expansion on a pure platinum surface in two orientations. The coverage agrees reasonably well with the experimental results. On these two surfaces there is no phase separation present, the coverages are smooth and have no abrupt drop as we have seen in our simulations.

6. Discussion and Perspectives

Although we have gained more insight into the surface coverage of RuO_2 and IrO_2 in chapter 3 we have also seen the shortcomings of the theoretical models today. In both phase diagrams 3.8.3 and 3.9.2 we can see at least one phase missing that is predicted by the theoretical one-electron processes but fails to be present in this approach. Possible reasons are the missing water layers that form the explicit solvent in the real system, missing description of ions in solution from the adjustment of pH levels and many other more intrinsic physical phenomena not included in the theoretical model. Method development to include those effects is in some cases ongoing by other groups but also in some cases not feasible at this point because of the computational costs that come with them. And for some there simply are no theoretical models available that provide a more accurate theoretical description besides explicit modeling, i.e. ions in solution that influence the pH level and through electrostatics also the surface adsorption and the capacitative layers on the surface of the catalyst. Nevertheless, this approach provides useful information about the change in surface states of both catalysts under investigation and from the point of computational costs it is cheap to produce.

A statistical method using probabilities of phase occurrence based on energy differences (Monte-Carlo based) approaches might be more suited as an approximation of the model to reach experimental results as was demonstrated in chapter 5. The underlying theory is

more sophisticated than in chapter 3 and computational costs are much higher. Still the fundamental shortcomings of the computational model are the same in the sense that it does not include and ion in solution, explicit solvents or capacitative effects. The major difference to the approach in chapter 3 is the inclusion of many possible surface structures and their statistical mixing based on the respective energies. This process allows to produce a more accurate description of the experimental result but also hints towards the missing theoretical descriptions as we have seen in the resulting computational CV in figure 5.10.3. The phase separation seems to be well covered by the currently available DFT models while the mixing of surface states during the potential increase between 0.5V and 1.2V is not well described. These two regions are fundamentally different as in the phase separation above 1.3V there are hydrogen atoms desorbing from the surface while in the mixing region there is a reordering of adsorbed hydrogen atoms. The desorption results in the breaking of bonds which is energetically well described in state-of-the-art DFT but the reordering does not include bond breaking but rather structural changes due to electrostatic effects resulting from the change in potential. As our theoretical model does not include these effects it is not surprising that those effects are missing in the resulting computational CV. With more sophisticated theoretical DFT implementations describing these effects one might be able to produce a more accurate computed CV from first principles.

The improvement of the oxygen evolution reaction in terms of energy efficiency is of interest when looking at large scale applications for energy storage or hydrogen production at industrial scales. A possible route for systematic improvement was demonstrated in chapter 4. Mixing other species into a catalyst during growth can result in an increase of activity. In this case mixing in between 11% and 14% results in a decrease of overpotential by about 0.2V in the experimental result. From the theoretical approach we see an improvement of 0.12V

which confirms the right trend in computations. The sensitivity of the method towards neighboring species and the coverage of the model slab with different adsorbates needs to be investigated further to ensure realistic descriptions of the system. This approach opens the door towards a systematic method to scan for possible new catalyst base materials that provide acceptable activities and stability in acidic solutions which can then be improved by mixing in other species during growth. While this procedure might be unfeasible experimentally due to the large number of possible combinations of species it is well accessible for theoretical methods with DFT today and provides the right trends in adsorption energetics.

As the theoretical models and the available implementations in current DFT codes are not fully descriptive of the complete phenomena that are at play in surface processes they can nevertheless provide useful insights to these processes and give hints towards the importance of specific effects. Today's DFT computations are thus not in themselves explaining complex material properties and processes in total but can suggest probable explanations about parts of the process and supply routes towards exploring the feasibility of new materials for a specific application. A continuous combined effort of experiment and theory promises the most efficient use of available resources and also the highest chance of success towards the design of new materials.

Bibliography

- [1] S. A. Montzka¹, E. J. Dlugokencky¹, and J. H. Butler. Non-co₂ greenhouse gases and climate change. *Nature Review*, 476, 2011.
- [2] Steven Chu and Arun Majumdar. Opportunities and challenges for a sustainable energy future. *Nature Perspective*, 488, 2012.
- [3] David C. Frank, Jan Esper, Christoph C. Raible, Ulf Büntgen, Valerie Trouet, Benjamin Stocker, and Fortunat Joos. Ensemble reconstruction constraints on the global carbon cycle sensitivity to climate. *Nature*, 463:527–530, 2010.
- [4] Climate change report card: These countries are reaching targets. National Geographic.
- [5] ‘we are in trouble.’ global carbon emissions reached a record high in 2018. The Washington Post.
- [6] Ioan Sarbu and Calin Sebarchievici. Chapter 4 - thermal energy storage. In Ioan Sarbu and Calin Sebarchievici, editors, *Solar Heating and Cooling Systems*, pages 99 – 138. Academic Press, 2017.
- [7] Shripad T. Revankar. Chapter six - chemical energy storage. In Hitesh Bindra and Shripad Revankar, editors, *Storage and Hybridization of Nuclear Energy*, pages 177 – 227. Academic Press, 2019.

- [8] Odne Stokke Burheim. Chapter 3 - mechanical energy storage. In Odne Stokke Burheim, editor, *Engineering Energy Storage*, pages 29 – 46. Academic Press, 2017.
- [9] Front matter. In Patrick T. Moseley and Jürgen Garche, editors, *Electrochemical Energy Storage for Renewable Sources and Grid Balancing*, page iii. Elsevier, Amsterdam, 2015.
- [10] The missing link: Why australia needs energy storage. <https://reneweconomy.com.au/>.
- [11] R de Levie. The electrolysis of water. *Journal of Electroanalytical Chemistry*, 476(1):92 – 93, 1999.
- [12] Roel van de Krol and Michael Grätzel. *Photoelectrochemical Hydrogen Production*. Springer US, 1 edition, 2012.
- [13] Zhi Wei Seh, Jakob Kibsgaard, Colin F. Dickens, Ib Chorkendorff, Jens K. Nørskov, and Thomas F. Jaramillo. Combining theory and experiment in electrocatalysis: Insights into materials design. *Science*, 355(6321), 2017.
- [14] Jesse D. Benck, Thomas R. Hellstern, Jakob Kibsgaard, Pongkarn Chakthranont, and Thomas F. Jaramillo. Catalyzing the hydrogen evolution reaction (her) with molybdenum sulfide nanomaterials. *ACS Catalysis*, 4(11):3957–3971, 2014.
- [15] Ding-Yuan Kuo, Jason K. Kawasaki, Jocienne N. Nelson, Jan Kloppenburg, Geoffroy Hautier, Kyle M. Shen, Darrell G. Schlom, and Jin Suntivich. Influence of surface adsorption on the oxygen evolution reaction on $\text{Fe}_2\text{O}_3(110)$. *Journal of the American Chemical Society*, 139(9):3473–3479, 2017. PMID: 28181433.
- [16] Daniel F. Abbott, Dmitry Lebedev, Kay Waltar, Mauro Povia, Maarten Nachtegaal, Emiliana Fabbri, Christophe Copéret,

- and Thomas J. Schmidt. Iridium oxide for the oxygen evolution reaction: Correlation between particle size, morphology, and the surface hydroxo layer from operando xas. *Chemistry of Materials*, 28(18):6591–6604, 2016.
- [17] J. K. Nørskov, T. Bligaard, J. Rossmeisl, and C. H. Christensen. Towards the computational design of solid catalysts. *Nature Chemistry*, 1(1):37–46, 2009.
- [18] T. Bligaard, J.K. Nørskov, S. Dahl, J. Matthiesen, C.H. Christensen, and J. Sehested. The brønsted–evans–polanyi relation and the volcano curve in heterogeneous catalysis. *Journal of Catalysis*, 224(1):206 – 217, 2004.
- [19] Francesc Viñes, Aleksandra Vojvodic, Frank Abild-Pedersen, and Francesc Illas. Brønsted–evans–polanyi relationship for transition metal carbide and transition metal oxide surfaces. *The Journal of Physical Chemistry C*, 117(8):4168–4171, 2013.
- [20] Isabela C. Man, Hai-Yan Su, Federico Calle-Vallejo, Heine A. Hansen, José I. Martínez, Nilay G. Inoglu, John Kitchin, Thomas F. Jaramillo, Jens K. Nørskov, and Jan Rossmeisl. Universality in oxygen evolution electrocatalysis on oxide surfaces. *ChemCatChem*, 3(7):1159–1165, 2011.
- [21] Youngmin Lee, Jin Suntivich, Kevin J. May, Erin E. Perry, and Yang Shao-Horn. Synthesis and activities of rutile FeO_2 and RuO_2 nanoparticles for oxygen evolution in acid and alkaline solutions. *The Journal of Physical Chemistry Letters*, 3(3):399–404, 2012. PMID: 26285858.
- [22] Gabor A. Somorjai and Jeong Y. Park. Molecular factors of catalytic selectivity. *Angewandte Chemie International Edition*, 47(48):9212–9228, 2008.

- [23] G.A. Somorjai and Minchul Yang. The surface science of catalytic selectivity. *Topics in Catalysis*, 24(1):61–72, Oct 2003.
- [24] K. Alavi. Molecular beam epitaxy. In K.H. Jürgen Buschow, Robert W. Cahn, Merton C. Flemings, Bernhard Ilshner, Edward J. Kramer, Subhash Mahajan, and Patrick Veyssi re, editors, *Encyclopedia of Materials: Science and Technology*, pages 5765 – 5780. Elsevier, Oxford, 2001.
- [25] Serhiy Cherevko, Simon Geiger, Olga Kasian, Nadiia Kulyk, Jan-Philipp Grote, Alan Savan, Buddha Ratna Shrestha, Sergiy Merzlikin, Benjamin Breitbach, Alfred Ludwig, and Karl J.J. Mayrhofer. Oxygen and hydrogen evolution reactions on ru, ruo2, ir, and iro2 thin film electrodes in acidic and alkaline electrolytes: A comparative study on activity and stability. *Catalysis Today*, 262:170 – 180, 2016. Electrocatalysis.
- [26] Allen J Bard and Larry R Faulkner. *Electrochemical Methods*. John Wiley & Sons, 2001.
- [27] Frank Marken, Andreas Neudeck, and Alan M. Bond. *Cyclic Voltammetry*, pages 51–97. Springer Berlin Heidelberg, Berlin, Heidelberg, 2002.
- [28] P. Hohenberg and W. Kohn. Inhomogeneous electron gas. *Phys. Rev.*, 136:B864–B871, Nov 1964.
- [29] W. Kohn and L. J. Sham. Self-consistent equations including exchange and correlation effects. *Phys. Rev.*, 140:A1133–A1138, Nov 1965.
- [30] D. Sholl and J.A. Steckel. *Density Functional Theory: A Practical Introduction*. Wiley, 2009.
- [31] N.W. Ashcroft and N.D. Mermin. *Solid State Physics*. Cengage Learning, 2011.

- [32] Fernando Nogueira, Miguel A L Marques, and Carlos Fiolhais. *A primer in density functional theory*. Lecture Notes in Physics. Springer, Berlin, 2003.
- [33] G. Kresse and J. Hafner. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B*, 47:558–561, Jan 1993.
- [34] G. Kresse and J. Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science*, 6(1):15–50, 1996.
- [35] Roger Fletcher. *Practical methods of optimization (2nd edition)*. Wiley, 1987.
- [36] Jorge Nocedal and S. Wright. *Numerical Optimization*. Springer-Verlag, NY, 2006.
- [37] F. Bloch. Bemerkungen zur Elektronentheorie des Ferromagnetismus und der elektrischen Leitfähigkeit. *Zeitschrift für Physik*, 57(7-8):545–555, 1929.
- [38] P. A. M. Dirac. Note on exchange phenomena in the Thomas atom. *Math. Proc. Cambridge Phil. Soc.*, 26(3):376–385, 1930.
- [39] E. Wigner. Effects of the electron interaction on the energy levels of electrons in metals. *Trans. Faraday Soc.*, 34:678–685, 1938.
- [40] J. P. Perdew and A. Zunger. Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B*, 23:5048–5079, May 1981.
- [41] J. P. Perdew and Y. Wang. Accurate and simple analytic representation of the electron-gas correlation energy. *Phys. Rev. B*, 45:13244–13249, Jun 1992.

- [42] U. von Barth and L. Hedin. A local exchange-correlation potential for the spin polarized case. i. *J. Phys. C: Solid State Physics*, 5(13):1629, 1972.
- [43] D. C. Langreth and M. J. Mehl. Beyond the local-density approximation in calculations of ground-state electronic properties. *Phys. Rev. B*, 28:1809–1834, Aug 1983.
- [44] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77:3865–68, 1996.
- [45] J. P. Perdew, K. Burke, and M. Ernzerhof. Errata: Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 78:1396, 1997.
- [46] B. Hammer, L. B. Hansen, and J. K. Nørskov. Improved adsorption energetics within density-functional theory using revised perdew-burke-ernzerhof functionals. *Phys. Rev. B*, 59:7413–7421, Mar 1999.
- [47] A. D. Becke. Density-functional exchange energy approximation with correct asymptotic behavior. *Phys. Rev. A*, 38(6):3098–3100, 1988.
- [48] C. Lee, W. Yang, and R. G. Parr. Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B*, 37(2):785–789, 1988.
- [49] N. E. Schultz, Y. Zhao, and D. G. Truhlar. Density functionals for inorganometallic and organometallic chemistry. *J. Phys. Chem. A*, 109(49):11127–11143, 2005.
- [50] J. Heyd, G. E. Scuseria, and M. Ernzerhof. Hybrid functionals based on a screened coulomb potential. *J. Chem. Phys.*, 124(219906), 2006.

- [51] J. P. Perdew, K. Burke, and M. Ernzerhof. Rationale for mixing exact exchange with density functional approximations. *J. Chem. Phys.*, 105(22):9982–9986, 1996.
- [52] M. Ernzerhof and J. P. Perdew. Generalized gradient approximation to the angle- and system-averaged exchange hole. *J. Chem. Phys.*, 109(3313), 1998.
- [53] R. J. Stephens, F. J. Devlin, C. F. Chabalowski, and M. J. Frisch. Ab initio calculation of vibrational absorption and circular dichroism spectra using density functional force fields. *J. Phys. Chem.*, 98(45):11623–11627, 1994.
- [54] I-Chun Lin, Ari P. Seitsonen, Maurício D. Coutinho-Neto, Ivano Tavernelli, and Ursula Rothlisberger. Importance of van der waals interactions in liquid water. *The Journal of Physical Chemistry B*, 113(4):1127–1131, 2009. PMID: 19123911.
- [55] Tobias Morawietz, Andreas Singraber, Christoph Dellago, and Jörg Behler. How van der waals interactions determine the unique properties of water. *Proceedings of the National Academy of Sciences*, 113(30):8368–8373, 2016.
- [56] M. Dion, H. Rydberg, E. Schröder, D. C. Langreth, and B. I. Lundqvist. Van der waals density functional for general geometries. *Phys. Rev. Lett.*, 92:246401, Jun 2004.
- [57] Kyuho Lee, Éamonn D. Murray, Lingzhu Kong, Bengt I. Lundqvist, and David C. Langreth. Higher-accuracy van der waals density functional. *Phys. Rev. B*, 82:081101, Aug 2010.
- [58] Alexandre Tkatchenko and Matthias Scheffler. Accurate molecular van der waals interactions from ground-state electron density and free-atom reference data. *Phys. Rev. Lett.*, 102:073005, Feb 2009.

- [59] Kiran Mathew, Ravishankar Sundararaman, Kendra Letchworth-Weaver, T. A. Arias, and Richard G. Hennig. Implicit solvation model for density-functional study of nanocrystal surfaces and reaction pathways. *The Journal of Chemical Physics*, 140(8):084106, 2014.
- [60] Kiran Mathew and Richard G. Hennig. Implicit self-consistent description of electrolyte in plane-wave density-functional theory. *arXiv e-prints*, page arXiv:1601.03346, Jan 2016.
- [61] Deniz Gunceler, Kendra Letchworth-Weaver, Ravishankar Sundararaman, Kathleen A Schwarz, and T A Arias. The importance of nonlinear fluid response in joint density-functional theory studies of battery systems. *Modelling and Simulation in Materials Science and Engineering*, 21(7):074005, oct 2013.
- [62] Ravishankar Sundararaman and William A. Goddard. The charge-asymmetric nonlocally determined local-electric (candle) solvation model. *The Journal of Chemical Physics*, 142(6):064107, 2015.
- [63] Oliviero Andreussi, Ismaila Dabo, and Nicola Marzari. Revised self-consistent continuum solvation in electronic-structure calculations. *The Journal of Chemical Physics*, 136(6):064102, 2012.
- [64] Ravishankar Sundararaman, Kathleen A. Schwarz, Kendra Letchworth-Weaver, and T. A. Arias. Spicing up continuum solvation models with salsa: The spherically averaged liquid susceptibility ansatz. *The Journal of Chemical Physics*, 142(5):054102, 2015.
- [65] Ravishankar Sundararaman, Kendra Letchworth-Weaver, and T. A. Arias. A computationally efficacious free-energy functional for studies of inhomogeneous liquid water. *The Journal of Chemical Physics*, 137(4):044107, 2012.

- [66] Ravishankar Sundararaman and T.A. Arias. Efficient classical density-functional theories of rigid-molecular fluids and a simplified free energy functional for liquid water. *Computer Physics Communications*, 185(3):818–825, 2014.
- [67] J. Rossmeisl, Z.-W. Qu, H. Zhu, G.-J. Kroes, and J.K. Nørskov. Electrolysis of water on oxide surfaces. *Journal of Electroanalytical Chemistry*, 607(1):83–89, 2007. Theoretical and Computational Electrochemistry.
- [68] Jörg Neugebauer and Matthias Scheffler. Adsorbate-substrate and adsorbate-adsorbate interactions of na and k adlayers on al(111). *Phys. Rev. B*, 46:16067–16080, Dec 1992.
- [69] V. Blum, R. Gehrke, F. Hanke, P. Havu, V. Havu, X. Ren, K. Reuter, and M. Scheffler. Ab initio molecular simulations with numeric atom-centered orbitals: FHI-AIMS. *Comp. Phys. Commun.*, 180:2175–2196, 2009.
- [70] X. Ren, P. Rinke, V. Blum, J. Wieferink, A. Tkatchenko, A. Sanfilippo, K. Reuter, and M. Scheffler. Resolution-of-Identity approach to Hartree-Fock, hybrid density functionals, RPA, MP2 and GW with numeric atom-centered orbital basis functions. *New J. Phys.*, 14(5):053020, may 2012.
- [71] Victor Wen zhe Yu, Fabiano Corsetti, Alberto García, William P. Huhn, Mathias Jacquelin, Weile Jia, Björn Lange, Lin Lin, Jianfeng Lu, Wenhui Mi, Ali Seifitokaldani, Álvaro Vázquez-Mayagoitia, Chao Yang, Haizhao Yang, and Volker Blum. Elsi: A unified software interface for kohn–sham electronic structure solvers. *Computer Physics Communications*, 222:267–285, 2018.
- [72] G. Kresse and J. Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B*, 54:11169–11186, Oct 1996.

- [73] G. Kresse and D. Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B*, 59:1758–1775, Jan 1999.
- [74] J. H. Ehrenreich Bohm, F. Seitz, and D. Turnbull (Herausgeber). Solid state physics, vol. 25, 420 seiten. academic press inc. publ., new york und london 1970. *Kristall und Technik*, 7(3):13–15, 1972.
- [75] CASM code. The CASM website.
<https://labs.materials.ucsb.edu/vandervan/anton/research/casm-code>.
- [76] J.M. Sanchez, F. Ducastelle, and D. Gratias. Generalized cluster description of multicomponent systems. *Physica A: Statistical Mechanics and its Applications*, 128(1):334–350, 1984.
- [77] J. M. Sanchez. Foundations and Practical Implementations of the Cluster Expansion. *Journal of Phase Equilibria and Diffusion*, 38(3):238–251, 2017.
- [78] Casm, v0.2.1 (2017). available from
<https://github.com/prisms-center/casmcode>. <https://github.com/prisms-center/CASMcode>. Accessed: 2019-07-31.
- [79] John C. Thomas and Anton Van der Ven. Finite-temperature properties of strongly anharmonic and mechanically unstable crystal phases from first principles. *Phys. Rev. B*, 88:214111, Dec 2013.
- [80] B. Puchala and A. Van der Ven. Thermodynamics of the zr-o system from first-principles calculations. *Phys. Rev. B*, 88:094108, Sep 2013.

- [81] A. Van der Ven, J.C. Thomas, Qingchuan Xu, and J. Bhattacharya. Linking the electronic structure of solids to their thermodynamic and kinetic properties. *Mathematics and Computers in Simulation*, 80(7):1393 – 1410, 2010. Multiscale modeling of moving interfaces in materials.
- [82] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning*. MIT press, 2006.
- [83] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT press, 2016.
- [84] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.
- [85] Manuel Calzolari. manuel-calzolari/sklearn-genetic: sklearn-genetic 0.2, apr 2019.
- [86] Volker Blum, Gus L. W. Hart, Michael J. Walorski, and Alex Zunger. Using genetic algorithms to map first-principles results to model hamiltonians: Application to the generalized ising model for alloys. *Phys. Rev. B*, 72:165113, Oct 2005.
- [87] Isabelle Guyon, Jason Weston, Stephen Barnhill, and Vladimir Vapnik. Gene selection for cancer classification using support vector machines. *Machine Learning*, 46(1):389–422, Jan 2002.
- [88] A. van de Walle and G. Ceder. Automating first-principles phase diagram calculations. *Journal of Phase Equilibria*, 23(4):348, Aug 2002.

- [89] Gus L. W. Hart, Volker Blum, Michael J. Walorski, and Alex Zunger. Evolutionary approach for determining first-principles hamiltonians. *Nature Materials*, 4:391–394, May 2005.
- [90] Nicholas Metropolis, Arianna W. Rosenbluth, Marshall N. Rosenbluth, Augusta H. Teller, and Edward Teller. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(6), 1953.
- [91] A van de Walle and M Asta. Self-driven lattice-model monte carlo simulations of alloy thermodynamic properties and phase diagrams. *Modelling and Simulation in Materials Science and Engineering*, 10(5):521–538, jul 2002.
- [92] Ding-Yuan Kuo, Hanjong Paik, Jan Kloppenburg, Brendan Faeth, Kyle M. Shen, Darrell G. Schlom, Geoffroy Hautier, and Jin Suntivich. Measurements of oxygen electroadsorption energies and oxygen evolution reaction on ruo₂(110): A discussion of the sabatier principle and its role in electrocatalysis. *Journal of the American Chemical Society*, 140(50):17597–17605, 2018.
- [93] Raúl J. Martín-Palma and Akhlesh Lakhtakia. Chapter 15 - vapor-deposition techniques. In Akhlesh Lakhtakia and Raúl J. Martín-Palma, editors, *Engineered Biomimicry*, pages 383 – 398. Elsevier, Boston, 2013.
- [94] Darrell schlom group. <http://schlom.mse.cornell.edu/>. Accessed: 31 July 2019.
- [95] Mohamed Henini. *Molecular Beam Epitaxy - From Research to Mass Production*. Elsevier, 1st edition, 2012.
- [96] Secondo Franchi. Chapter 1 - molecular beam epitaxy: fundamentals, historical background and future prospects. In Mohamed Henini, editor, *Molecular Beam Epitaxy*, pages 1–46. Elsevier, Oxford, 2013.

- [97] J. K. Nørskov, J. Rossmeisl, A. Logadottir, L. Lindqvist, J. R. Kitchin, T. Bligaard, and H. Jónsson. Origin of the overpotential for oxygen reduction at a fuel-cell cathode. *The Journal of Physical Chemistry B*, 108(46):17886–17892, 2004.
- [98] Karsten Reuter and Matthias Scheffler. Composition, structure, and stability of $\text{RuO}_2(110)$ as a function of oxygen pressure. *Phys. Rev. B*, 65:035406, Dec 2001.
- [99] Vasp wiki website. https://cms.mpi.univie.ac.at/wiki/index.php/Vibrational_frequencies_of_CO_on_Ni_111_surface. Accessed: 10. Sep 2019.
- [100] US Government. National institute of standards and technology. <http://www.nist.gov>, 2012. US Department of Commerce.
- [101] Nist website. <https://www.nist.gov/pml/productsservices/physical-reference-data>. Accessed: 10. Sep 2019.
- [102] Michael J. Gillan, Dario Alfè, and Angelos Michaelides. Perspective: How good is dft for water? *The Journal of Chemical Physics*, 144(13):130901, 2016.
- [103] Rentao Mu, David C. Cantu, Xiao Lin, Vassiliki-Alexandra Glezakou, Zhitao Wang, Igor Lyubinetsky, Roger Rousseau, and Zdenek Dohnálek. Dimerization induced deprotonation of water on $\text{RuO}_2(110)$. *The Journal of Physical Chemistry Letters*, 5(19):3445–3450, 2014. PMID: 26278461.
- [104] Rentao Mu, David C. Cantu, Vassiliki-Alexandra Glezakou, Igor Lyubinetsky, Roger Rousseau, and Zdenek Dohnálek. Deprotonated water dimers: The building blocks of segmented water chains on rutile $\text{RuO}_2(110)$. *The Journal of Physical Chemistry C*, 119(41):23552–23558, 2015.

- [105] F. Abild-Pedersen, J. Greeley, F. Studt, J. Rossmeisl, T. R. Munter, P. G. Moses, E. Skúlason, T. Bligaard, and J. K. Nørskov. Scaling properties of adsorption energies for hydrogen-containing molecules on transition-metal surfaces. *Phys. Rev. Lett.*, 99:016105, Jul 2007.
- [106] Sarah Gautier, Stephan N. Steinmann, Carine Michel, Paul Fleurat-Lessard, and Philippe Sautet. Molecular adsorption at pt(111). how accurate are dft functionals? *Phys. Chem. Chem. Phys.*, 17:28921–28930, 2015.
- [107] Katrin Tonigold and Axel Groß. Adsorption of small aromatic molecules on the (111) surfaces of noble metals: A density functional theory study with semiempirical corrections for dispersion effects. *The Journal of Chemical Physics*, 132(22):224701, 2010.
- [108] Ravishankar Sundararaman, Kendra Letchworth-Weaver, Kathleen A. Schwarz, Deniz Gunceler, Yalcin Ozhables, and T.A. Arias. Jdftx: Software for joint density-functional theory. *SoftwareX*, 6:278–284, 2017.
- [109] Eva M. Fernández, Poul G. Moses, Anja Toftelund, Heine A. Hansen, José I. Martínez, Frank Abild-Pedersen, Jesper Kleis, Berit Hinnemann, Jan Rossmeisl, Thomas Bligaard, and Jens K. Nørskov. Scaling relationships for adsorption energies on transition metal oxide, sulfide, and nitride surfaces. *Angewandte Chemie International Edition*, 47(25):4683–4686, 2008.
- [110] Niels Bendtsen Halck, Valery Petrykin, Petr Krtil, and Jan Rossmeisl. Beyond the volcano limitations in electrocatalysis – oxygen evolution reaction. *Phys. Chem. Chem. Phys.*, 16:13682–13688, 2014.

- [111] G.L. Elizarova, G.M. Zhidomirov, and V.N. Parmon. Hydroxides of transition metals as artificial catalysts for oxidation of water to dioxygen. *Catalysis Today*, 58(2):71 – 88, 2000.
- [112] Yang Yang, Lianna Dang, Melinda J. Shearer, Hongyuan Sheng, Wenjie Li, Jie Chen, Peng Xiao, Yunhuai Zhang, Robert J. Hamers, and Song Jin. Highly active trimetallic nifecr layered double hydroxide electrocatalysts for oxygen evolution reaction. *Advanced Energy Materials*, 8(15):1703189, 2018.
- [113] Runbang Tang, Yuefeng Nie, Jason K. Kawasaki, Ding-Yuan Kuo, Guido Petretto, Geoffroy Hautier, Gian-Marco Rignanese, Kyle M. Shen, Darrell G. Schlom, and Jin Suntivich. Oxygen evolution reaction electrocatalysis on sriro3 grown using molecular beam epitaxy. *J. Mater. Chem. A*, 4:6831–6836, 2016.
- [114] Colin F. Dickens, Charlotte Kirk, and Jens K. Nørskov. Insights into the electrochemical oxygen evolution reaction with ab initio calculations and microkinetic modeling: Beyond the limiting potential volcano. *The Journal of Physical Chemistry C*, 123(31):18960–18977, 2019.
- [115] A. van de Walle, M. D. Asta, and G. Ceder. The Alloy Theoretic Automated Toolkit: A user guide. *Calphad*, 26:539–553, 2002.
- [116] A. van de Walle. Multicomponent multisublattice alloys, non-configurational entropy and other additions to the Alloy Theoretic Automated Toolkit. *Calphad*, 33:266–278, 2009.
- [117] Eriko Watanabe, Jan Rossmeisl, Mårten E. Björketun, Hiroshi Ushiyama, and Koichi Yamashita. Atomic-scale analysis of the ruo2/water interface under electrochemical conditions. *The Journal of Physical Chemistry C*, 120(15):8096–8103, 2016.

- [118] Ian T. McCrum and Michael J. Janik. ph and alkali cation effects on the pt cyclic voltammogram explained using density functional theory. *The Journal of Physical Chemistry C*, 120(1):457–471, 2016.
- [119] G. S. Karlberg, T. F. Jaramillo, E. Skúlason, J. Rossmeisl, T. Bligaard, and J. K. Nørskov. Cyclic voltammograms for h on pt(111) and pt(100) from first principles. *Phys. Rev. Lett.*, 99:126101, Sep 2007.
- [120] M. Scheffler and V. Blum. Theoretical Materials Science. Lecture Notes, 2008.
- [121] M. Born and R. Oppenheimer. Zur quantentheorie der molekeln. *Annalen der Physik*, 389(20):457–484, 1927.
- [122] H. A. Jahn, E. Teller, and Frederick George Donnan. Stability of polyatomic molecules in degenerate electronic states – l-orbital degeneracy. *Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences*, 161(905):220–235, 1937.
- [123] P. Pulay. Convergence acceleration of iterative sequences. the case of scf iteration. *Chem. Phys. Lett.*, 73(2):393 – 398, 1980.

Appendix

I. DFT equations and background of methods

This chapter aims to very briefly show the main foundations DFT is built on. It is in close relation to many textbooks [30, 31, 32], to cite a few. The particular introduction shown here is compiled from my own lecture notes [120].

For quantum mechanical solid state theory the central equation to solve is the many-body Schrödinger equation:

$$[\hat{T}^e + \hat{T}^{\text{ion}} + \hat{U}^{\text{e-e}} + \hat{U}^{\text{e-ion}} + \hat{U}^{\text{ion-ion}}] \Psi = E\Psi \quad (\text{I.1})$$

wherein $\Psi \equiv \Psi(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{R}_1, \dots, \mathbf{R}_M)$ is the all-electron all-nuclei wave function. Further, the terms of kinetic energy T have electronic contributions described by

$$\hat{T}^e = \sum_{k=1}^N \frac{\mathbf{p}_k^2}{2m}, \quad (\text{I.2})$$

and nuclear contributions captured by

$$\hat{T}^{\text{ion}} = \sum_{l=1}^M \frac{\mathbf{P}_l^2}{2M_l}. \quad (\text{I.3})$$

Potential energy contributions are given by three possible interactions, the electron-electron interaction given by

$$\hat{U}^{\text{e-e}} = \frac{1}{2} \frac{1}{4\pi\epsilon_0} \sum_{k \neq k'}^{N,N} \frac{e^2}{|\mathbf{r}_k - \mathbf{r}_{k'}|}, \quad (\text{I.4})$$

the electron-ion interactions described by

$$\hat{U}^{\text{e-ion}} = \frac{1}{2} \frac{1}{4\pi\epsilon_0} \sum_{l \neq l'}^{M,M} \frac{e^2}{|\mathbf{R}_l - \mathbf{R}_{l'}|}, \quad (\text{I.5})$$

and the ion-ion interaction from

$$\hat{U}^{\text{ion-ion}} = \sum_{k=1}^N \sum_{l=1}^M \mathbf{v}_l^{\text{ion}} (|\mathbf{R}_l - \mathbf{r}_k|) . \quad (\text{I.6})$$

With these we have a complete description of the electronic structure problem without having to apply any approximations. The main idea of Kohn and Sham was to reduce the mathematical problem size from Ψ containing $\sim 10^{23}$ electronic coordinates to the electronic density $n(\mathbf{r})$ having only three spatial coordinates. The proof of the existence and uniqueness is given in [29].

I.1. Born-Oppenheimer Approximation

For practical mathematical reasons it is favorable to decouple the movement of the ion cores from the movement of the electrons. This is the Born-Oppenheimer approximation [121]. Comparing the mass of an electron $m_e = 9.11 \cdot 10^{-31}$ kg the mass of the lightest nucleus, the proton with $m_p = 1.67 \cdot 10^{-27}$ kg we arrive at $m_p/m_e \approx 1836$ meaning the proton is roughly 2000 times heavier than the electron. From a physical point of view the electron moves too fast for the proton to follow and thus a decoupling of both is well justified. Taking into account the most of the time much heavier elements than hydrogen are described by DFT this argumentation solidifies. Mathematically one writes

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_N, \mathbf{R}_1, \dots, \mathbf{R}_M) = \sum_v \Lambda_v(\{\mathbf{R}_m\}) \Phi_v\{\mathbf{R}_m\}(\{\mathbf{r}_k\}) \quad (\text{I.7})$$

which allows to formally implement the DFT equations in computer code without the dependency on nuclear coordinates $\{\mathbf{R}_m\}$. However, there are known problems with this approximation and when one wants to look at properties connected to or arising from

- Polaron-induced superconductivity

- Dynamical Jahn-Teller effect[122]
- Diffusion phenomena
- Non-adiabatic processes and chemical reactions
- Charge carriers / transport effects

Special care has to be taken of the coupling between electrons and nuclei. None of these mentioned effects will be investigated and the Born-Oppenheimer approximation will be employed during this work.

I.2. SCF method

To solve the electronic problem we need to use a guess at the start because the result we are looking for depends on the density we seek. We thus need to supply an initial guess for the density $n(\mathbf{r})$ which then gets successively improved by mixing algorithms [123] and solving the electronic equations. This procedure is illustrated in figure I.1 and described in detail in [31, 69]. This procedure is in principle identical for all DFT codes and follows that yellow loop in figure I.1. In FHI-aims there is additional infrastructure marked in red boxes that improve computational performance only and is detailed in [71].

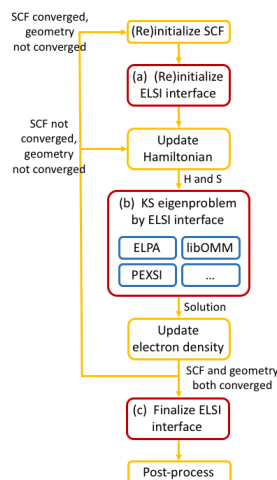


Figure I.1.: The SCF cycle to obtain a converged density $n(\mathbf{r})$ in FHI-aims[71].

II. CASM: Command line output

- (a) initialize the project by
`casm init`
- (b) Set a composition axes by: `casm composition -s N` where N is the axis number displayed from the code
- (c) Choose the minimum and maximum cell volume up to which the expansion shall generate unrelaxed structures (this can be changed later if the input data is not sufficient) `casm enum -max 5 -m 2` to set maximum volume to 5
- (d) Direct enumeration of all possible structures in the chosen volume `casm enum -a -m 0`
- (e) Select subsets of structures for computation: `casm select -set 'le(scel_size,2)'` selects all structures with super cell volume (scel_size) lower or equal (le) than 2.
- (f) Submit computations directly into the compute queue: `casm-calc -submit`
- (g) After computations are done read information: `casm update`

A query to the CASM code about the current status of the project can be requested by `casm status -d` which shows for this project:

```
1) Project initialized: TRUE

- Project name: IrO2
- Project location: /home/jank/catalys/RFE_fitting
- Lattice point group size: 8
- Lattice point group is D2h
- Factor group size: 1
- Crystal point group is: C1
```


2) Composition axes

- Composition axes selected: TRUE

3) Generate configurations

- Number of supercells generated: 11

- Number of configurations generated: 212

- Number of configurations currently selected: 4

4) Calculate configuration properties

- Current calctype: default

- Current cluster expansion: formation_energy

- Number of configurations calculated: 212 / 212 generated

INDEX	SUPERCELL	#CONFIGS G / C / S
0	SCEL1_1_1_1_0_0_0	6 / 6 / 0
1	SCEL2_2_1_1_0_0_0	18 / 18 / 0
2	SCEL2_1_2_1_0_0_0	17 / 17 / 0
3	SCEL2_2_1_1_0_0_1	17 / 17 / 0
4	SCEL3_3_1_1_0_0_1	22 / 22 / 0
5	SCEL3_1_3_1_0_0_0	13 / 13 / 0
6	SCEL3_3_1_1_0_0_2	21 / 21 / 0
7	SCEL3_3_1_1_0_0_0	16 / 16 / 0
8	SCEL4_2_2_1_0_0_0	26 / 26 / 2
9	SCEL4_4_1_1_0_0_0	35 / 35 / 0
10	SCEL5_5_1_1_0_0_0	21 / 21 / 2

TOTAL 212 / 212 / 4

G:Generated, C:Calculated, S:Selected

5) Choose chemical reference

- Chemical reference set: TRUE

To show the chemical reference, run `'casm ref -d'`

6) Generate basis functions: TRUE

7) Fit effective cluster interactions (ECI): TRUE

It shows a summary of what the project contains. In section 4) we see a detailed list of structures in the project. The name of the super cell is created by the following rule:

- (a) SCEL: short for super cell
- (b) First number after SCEL: Cell volume calculated by $i \cdot j \cdot k$ with i , j and k as the multiplication factors of the primitive input cell
- (c) Following three numbers: Multiplication factors of the primitive input cell used to generate the super cell (directions of (a,b,c) lattice vectors)
- (d) Last three numbers: Super cell transformation factors (i.e. from a tetragonal to a hexagonal cell)

For this project the super cell generator was instructed to multiply in the plane of the (a,b)-lattice vectors only since we have a surface as primitive input that has the vacuum vector in c-direction. Once these steps are done one ends up with DFT input data for the cluster expansion. The decision of whether the computed input data is sufficiently accurate needs to be made carefully. Simply only from the number of data points this decision can not be made. The DFT input data for the cluster expansion of this project is shown in figure 5.6.1. Before going further into the cluster expansion we can perform analysis similar to the previous chapter by looking at surface phase diagrams of all structures that we have computed.

III. Monte-Carlo settings:

Shown here is a single input file for the Monte-Carlo simulation. Parameters are adjusted as stated in the text for different runs, i.e. super cell size.

```
{
  "comment" : "Attributes prepended with '_' are ignored.",
  "debug" : false,
  "ensemble" : "grand_canonical",
  "method" : "metropolis",
  "model" : {
    "formation_energy" : "formation_energy"
  },
  "supercell" : [
    [60, 0, 0],
    [0, 60, 0],
    [0, 0, 1]
  ],
  "data" : {
    "sample_by" : "pass",
    "sample_period" : 10,
    "_N_sample" : 1000,
    "_N_pass" : 1000,
    "_N_step" : 1000,
    "_max_pass" : 10000,
    "min_pass" : 1000,
    "_max_step" : 10000,
    "_max_sample" : 500,
    "_min_sample" : 100,
    "confidence" : 0.99,
    "measurements" : [
      {
        "quantity" : "formation_energy",
```

```
    "precision" : 1e-3
  },
  {
    "quantity" : "potential_energy"
  },
  {
    "quantity" : "atom_frac"
  },
  {
    "quantity" : "site_frac"
  },
  {
    "quantity" : "comp",
    "precision" : 1e-3
  },
  {
    "quantity" : "comp_n"
  },
  {
    "quantity" : "all_correlations"
  }
],
"storage" : {
  "write_observations" : true,
  "write_trajectory" : true,
  "output_format" : ["json"]
},
"_enumeration": {
  "check" : "eq(1,1)",
  "metric" : "clex_hull_dist(ALL)",
  "insert_canonical" : true,
  "check_existence" : true,
  "N_halloffame" : 100,
```

```
        "sample_mode" : "on_sample"
    }
},
"driver" : {
    "mode" : "incremental",
    "dependent_runs" : true,
    "motif" : {
        "configname" : "auto",
        "_configname" : "SCEL3_3_1_1_0_2_2/0",
        "_configdof" : "path/to/final_state.json"
    },
    "initial_conditions" : {
        "param_chem_pot" : {
            "a" : -4.00
        },
        "temperature" : 300.0,
        "tolerance" : 0.01
    },
    "final_conditions" : {
        "param_chem_pot" : {
            "a" : 7.0
        },
        "temperature" : 300.0,
        "tolerance" : 0.01
    },
    "incremental_conditions" : {
        "param_chem_pot" : {
            "a" : 0.05
        },
        "temperature" : 0.0,
        "tolerance" : 0.01
    }
}
```

}

IV. Basis set parameters and resulting fit for ZPE CE

```
{
  "basis_functions" : {
    "site_basis_functions" : "occupation"
  },
  "orbit_branch_specs" : {
    "2" : {"max_length" : 14.0},
    "3" : {"max_length" : 10.0},
    "4" : {"max_length" : 8.0}
  }
}
```

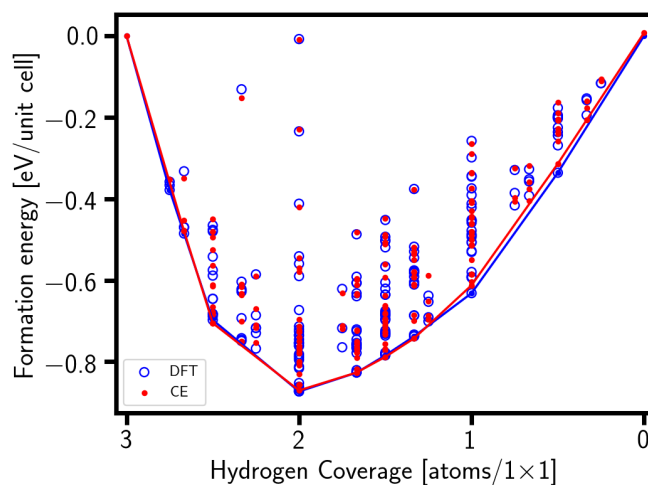


Figure IV.1.: Fitting result for ZPE added into CE

V. Ti-alloyed host structures

All tested positions for the titanium-alloyed IrO_2 :

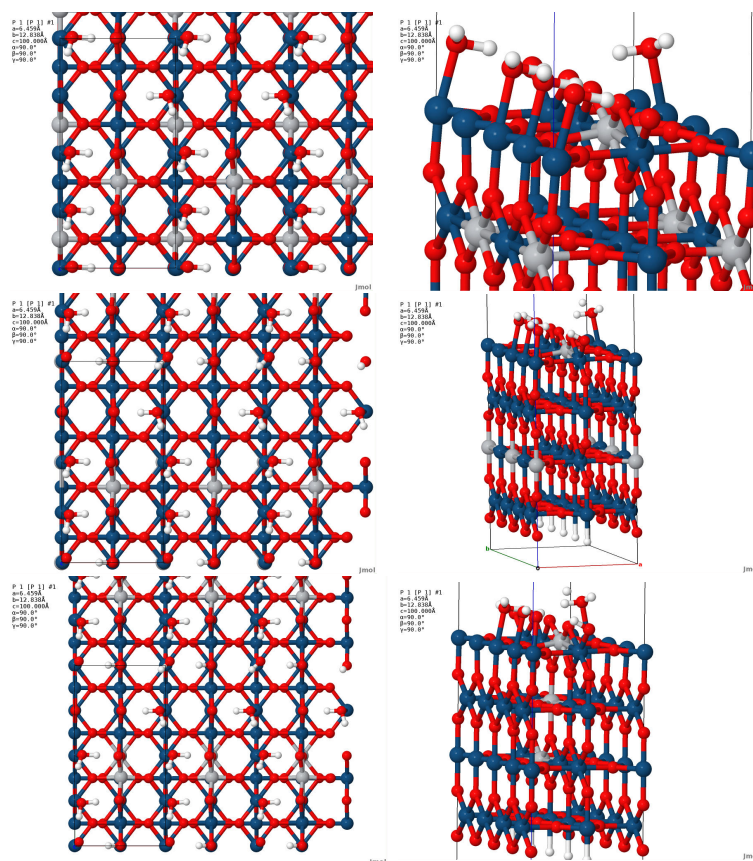


Figure V.1.: Structure 1 2 3

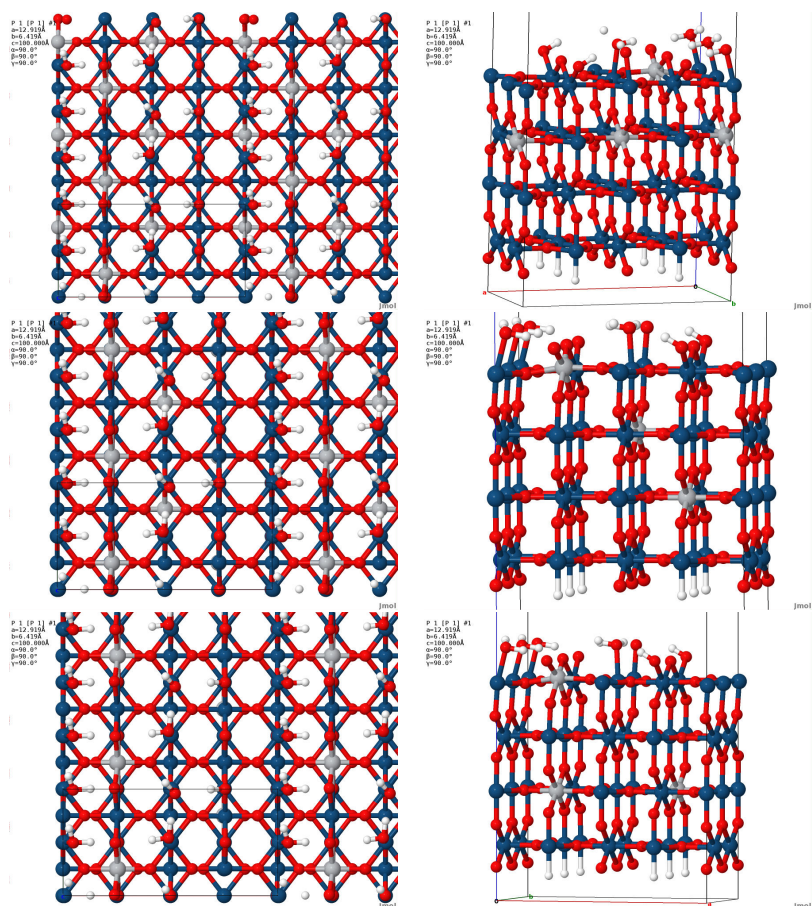


Figure V.2.: Structure a, b, c

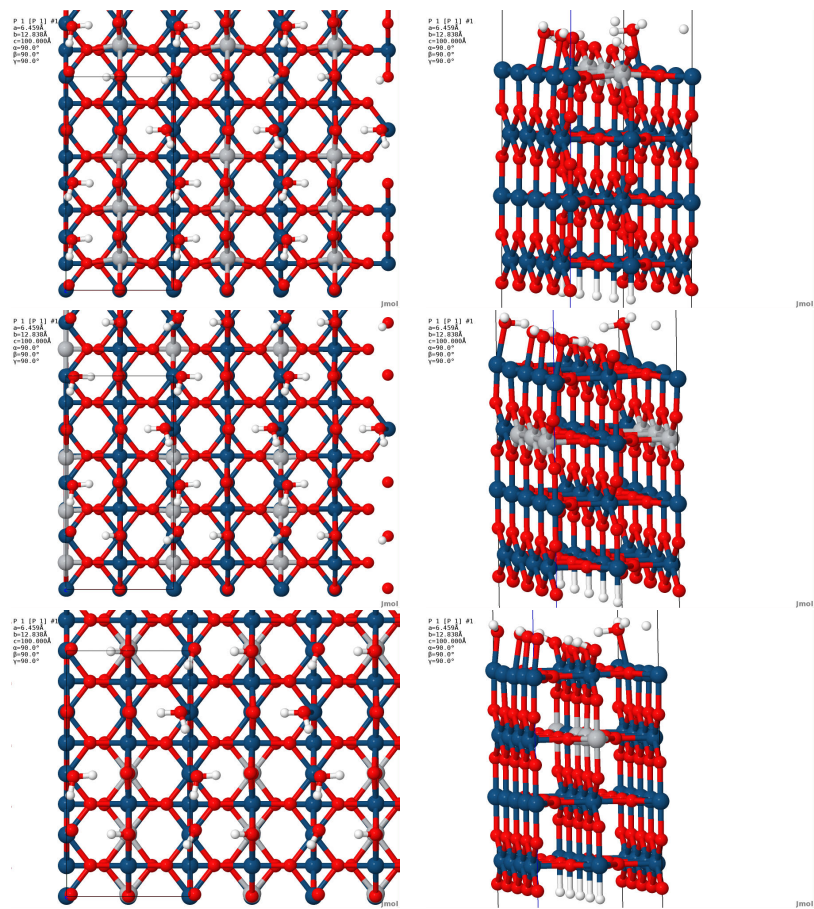


Figure V.3.: Structure p1, p2, p3

VI. Monte-Carlo Snapshots

Snapshots from the Monte-Carlo simulation

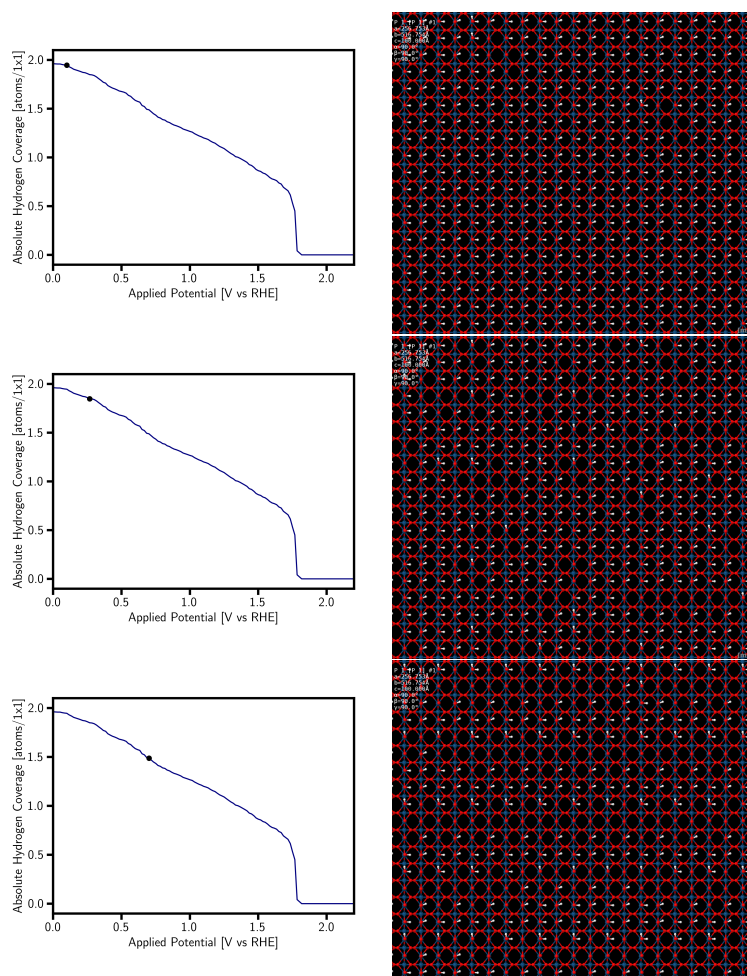


Figure VI.1.: MC snapshots

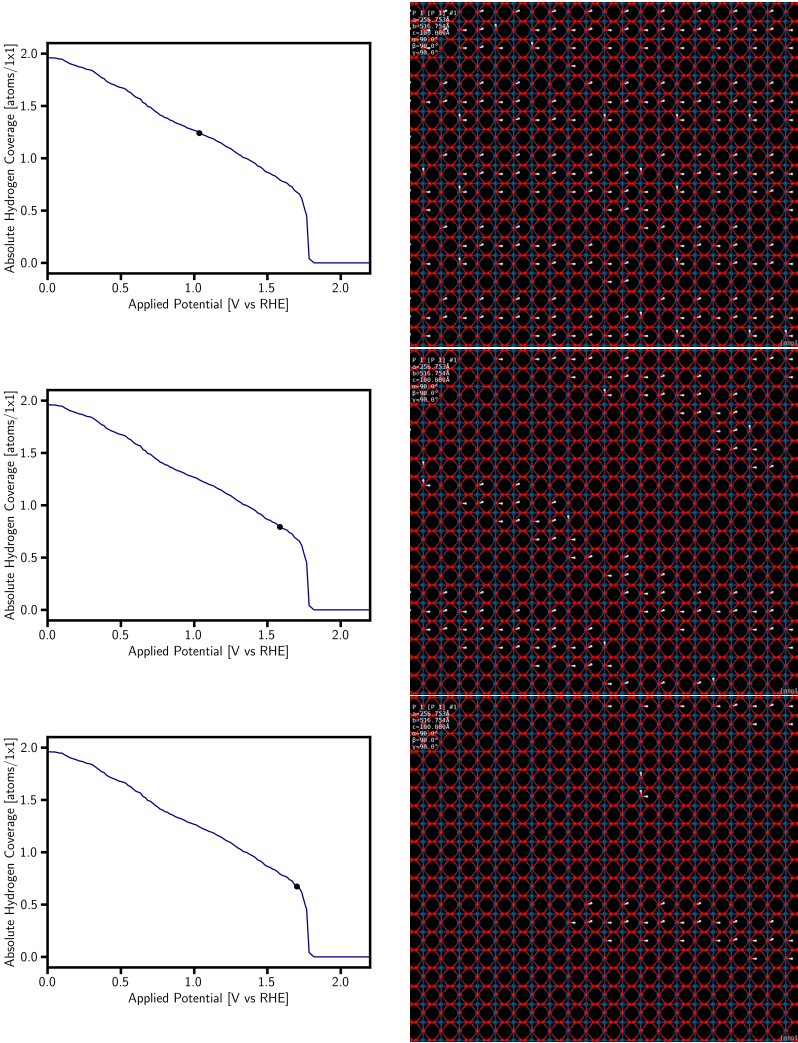
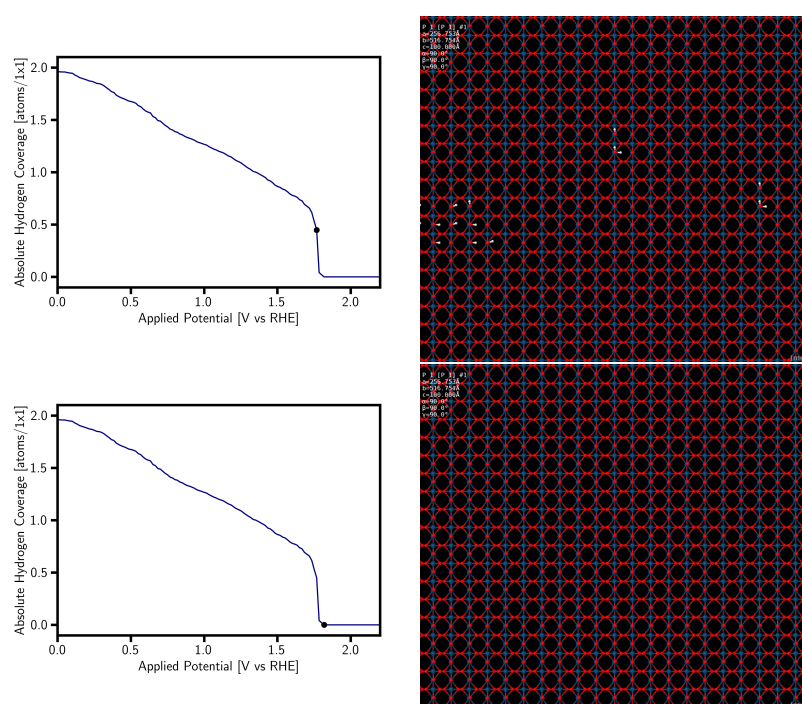


Figure VI.2.: MC snapshots

**Figure VI.3.: MC snapshots**