Dissertation

Atomistic Investigations on Plasma Facing Components of Beryllium and Tungsten

Lei Chen

Supervisor: Prof. Dr. Michael Probst
Co-Supervisor: Dr. Alexander Kaiser

Institut für Ionenphysik und Angewandte Physik
and
Doctoral Programme Computational Interdisciplinary Modelling (DK CIM)

University of Innsbruck

Innsbruck, 7 February 2020
Abstract

Beryllium (Be) and Tungsten (W) will be employed as the armor of the first wall and the divertor components, respectively, of the International Thermonuclear Experimental Reactor (ITER). The plasma confinement in a tokamak cannot be perfect. Plasma instabilities which include Edge-Localized Modes (ELM) occurring in the edge of the High-Confinement Mode (H-mode), Vertical Displacement Events (VDE) and disruptions pose a serious threat to the wall materials. The bombardment with hot plasma has various consequences on the Plasma Facing Materials (PFM), including sputtering, adsorption, absorption, vacancy formation, etc. which limits the lifetime of the components. Vice versa, the erosion materials enter into the plasma as impurities. On one side these impurities induce radiation losses in the plasma, decrease the plasma density and even potentially terminate the fusion reaction, while, on the other side they can redeposit on the PFM leading to the formation of mixed materials. Thus Plasma Wall Interaction (PWI) is a complex process that is and will continue to be a key interest in fusion research. Simulation experiments performed in tokamaks, in linear plasma devices, in electron and in ion beam load facilities provide key information for studying PWI. However, a comprehensive understanding at the atomistic level can only be achieved with a joint effort of experiments and computational modeling. In this thesis, I focused on the modeling at the atomistic level.

The main goal of the project was the development of machine learned potential energy surfaces for dynamic simulations of chemical reactions that occur on the surfaces in a tokamak and the application of these potentials in simulations. The underlying electronic structure theory is Density Functional Theory (DFT), due to the size and number of structures that have to be included in the training data for the potential energy fitting. Therefore, one of the early goals in my thesis was to benchmark density functionals for ground state electronic properties of a set of small neutral molecules containing Be, W, and H. A study of the basis set dependence of the couple cluster CCSD(T) energies showed that CCSD(T)/cc-pVQZ energies can be used as reference since they agree well with the Complete Basis Set (CBS) limit from a cc-pVQZ/cc-pV5Z extrapolation level and with available experimental data. All-electron calculations for molecules containing W indicated that core correlations contributed 3-5% for W hydrides and W$_2$ in atomization energies. The range-separated hybrid meta-GGA functional M11 on average shows accurate results for predicting atomization energies, bond lengths and vibrational frequencies for the studied molecules.

Successively, plane wave DFT was employed to calculate the surface defect energetics for pure Be and W and mixed materials (Be$_2$W and Be$_{12}$W). The energetics for different surface slabs indict that atomistic modeling is indeed essential for understanding the erosion and adsorption of wall materials in ITER. The plane wave DFT method was also used to investigate hydrogen adsorption and retention on the Be$_2$W surface. We found
a bridge site between two tungsten atoms to be the most stable adsorption site. The calculated diffusion barrier via this bridge site to a subsurface site is 1.53 eV, and the corresponding diffusion rate is $1.6 \times 10^{-6} \exp(-1.35 \text{ eV/kT}) \ \text{m}^2/\text{s}$.

For modeling of sputtering processes, dynamic simulations are necessary. Therefore, we have developed an interface between the neural network library aenet and the Molecular Dynamics (MD) code DL_POLY to be able to test the fitted Neural Network Potential (NNP) in dynamic simulations. However, only the total energy of configurations but no atomic forces are used to fit the NNPs in aenet 2.0.3. We found that force fitting is essential for developing NNPs applied in more complex MD simulations like the sputtering simulations in this thesis. Thus, we changed to use another code, n2p2 in which both energies and forces are used to train NNPs and applied it for sputtering simulations of Be surfaces.

Static plane wave DFT and ab initio molecular dynamic simulations were both used to generate and augment training data for the artificial neural networks. A neural network potential energy function for the beryllium surface for the beryllium self-sputtering process was such trained successfully by a refinement procedure. On this system we could demonstrate that the neural network potential can deliver ab initio MD sputtering yields in good agreement with previous simulations. In the last work included in this thesis, the sputtering of a beryllium surface by D impacts was investigated by MD simulations using such newly generated neural network potentials.
Acknowledgments

Firstly, I would like to express my sincere thanks to my supervisor Dr. Alexander Kaiser for his motivation, encouragement, broad knowledge and continuous support for my PhD research. I am very grateful to Prof. Dr. Michael Probst for his kindly help for applying scholarships and insightful comments on the publications. My sincere thanks also goes to other group members, especially to Daniel Süß, MSc., Priv.-Doz. Dr. Andreas Mauracher and Dr. Ivan Sukuba.

I want to thank my parents, my family, my boyfriend and my friends for supporting me spiritually throughout my PhD study and my life in Innsbruck.

I appreciate the University of Innsbruck for providing me such an advanced academic atmosphere. This work was supported by Austrian Science Fund (FWF) project under Grant No. P28979-N27 in the framework of the Doctoral Programme Computational Interdisciplinary Modelling (DK CIM). Part of this work was supported by the Kommission zur Koordination der Kernfusionsforschung der ÖAW (KKKÖ) Matching Grant, and by the Tyrolean Science Fund (TWF) Project UNI-0404/2308.

The work was also supported by the EUROFusion consortium via the Euratom research and training programme 2014-2018 and 2019-2020 under grant agreement No 633053.

Almost all the simulations in this work was carried out on the high-performance supercomputer, thus I want to thank the members who maintain LEO clusters and Vienna Scientific Cluster (VSC).
Contents

1 List of publications 4

2 Introduction 6
  2.1 Nuclear Fusion 6
  2.2 ITER project 9
  2.3 Plasma facing materials 10
  2.4 Plasma wall interaction 13
    2.4.1 Erosion 13
    2.4.2 Hydrogen isotopes in PFM 20

3 Modeling methods 24
  3.1 Quantum chemical calculations 24
    3.1.1 Hartree Fock method 24
    3.1.2 Complete basis set limit 26
    3.1.3 All-electron calculations versus effective core potentials 27
    3.1.4 Density functional theory 28
    3.1.5 Density functional theory with plane wave basis sets 30
  3.2 Classical molecular dynamic simulations 31
    3.2.1 Newton’s equation of motion 31
    3.2.2 Integrators for the equations of motion 32
    3.2.3 Thermodynamic Ensembles 33
    3.2.4 Force fields 34
    3.2.5 Ab initio molecular dynamic simulations 36
  3.3 Neural network approach 37
    3.3.1 Behler-Parrinello high dimensional feed forward neural network potential 40
    3.3.2 Iterative training set refinement 42
4 **Summary of results**

4.1 Performance of DFT functionals for neutral Be/W/H molecules .... 53
4.2 DFT study for surface defect energetics of Be, W and their alloys ... 56
4.3 H adsorption and penetration into the Be$_2$W surface .............. 57
  4.3.1 Details of the calculations ........................................ 57
  4.3.2 Surface model .................................................... 57
  4.3.3 Single hydrogen atom adsorption on slab D ....................... 58
  4.3.4 Single hydrogen atom adsorption on slab A ....................... 62
  4.3.5 Penetration of H atoms into the Be$_2$W subsurface ................ 63
4.4 Neural Network potential energy functions and their application to the self-sputtering of a Be surface .......................... 66
4.5 Development of NNP for Be sputtering simulation by D impacts .... 67
  4.5.1 Training data generation and refinement .......................... 67
  4.5.2 Non-cumulative sputtering simulations of a Be(0001) surface by D impacts using NNP .............................. 68

5 **Outlook on future research**

6 **Conferences**

**Bibliography**

**Appendix**

Article 1: Performance of DFT functionals for properties of small molecules containing beryllium, tungsten and hydrogen .................. 93
Article 2: Beryllium, tungsten and their alloys Be$_2$W and Be$_{12}$W: Surface defect energetics from density functional theory calculations ............... 101
Article 3: A neural network interface for DL_POLY and its application to liquid water ..................................................... 110
Article 4: Iterative training set refinement enables reactive molecular dynamics via machine learned forces .......................... 118
# Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ASDEX</td>
<td>Axially Symmetric Divertor Experiment</td>
</tr>
<tr>
<td>Be</td>
<td>Beryllium</td>
</tr>
<tr>
<td>BOMD</td>
<td>Born-Oppenheimer Molecular Dynamics</td>
</tr>
<tr>
<td>BOP</td>
<td>Bond Order Potential</td>
</tr>
<tr>
<td>CBS</td>
<td>Complete Basis Set</td>
</tr>
<tr>
<td>CD</td>
<td>Charge Density</td>
</tr>
<tr>
<td>CFC</td>
<td>Carbon Fibre Composite</td>
</tr>
<tr>
<td>CI</td>
<td>Configuration Interaction</td>
</tr>
<tr>
<td>DFT</td>
<td>Density Functional Theory</td>
</tr>
<tr>
<td>EAST</td>
<td>Experimental Advanced Superconducting Tokamak</td>
</tr>
<tr>
<td>ECP</td>
<td>Effective Core Potentials</td>
</tr>
<tr>
<td>ELM</td>
<td>Edge-Localized Modes</td>
</tr>
<tr>
<td>GGA</td>
<td>Generalized Gradient Approximation</td>
</tr>
<tr>
<td>GTO</td>
<td>Gaussian Type Orbital</td>
</tr>
<tr>
<td>H-mode</td>
<td>High-Confinement Mode</td>
</tr>
<tr>
<td>HF</td>
<td>Hartree-Fock</td>
</tr>
<tr>
<td>ITER</td>
<td>International Thermonuclear Experimental Reactor</td>
</tr>
<tr>
<td>JET</td>
<td>Joint European Torus</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>KS</td>
<td>Kohn-Sham</td>
</tr>
<tr>
<td>LDA</td>
<td>Local Density Approximation</td>
</tr>
<tr>
<td>MD</td>
<td>Molecular Dynamics</td>
</tr>
<tr>
<td>ML</td>
<td>Machine Learning</td>
</tr>
<tr>
<td>NN</td>
<td>Neural Network</td>
</tr>
<tr>
<td>NNP</td>
<td>Neural Network Potential</td>
</tr>
<tr>
<td>NVE</td>
<td>micro-canonical ensemble</td>
</tr>
<tr>
<td>NVT</td>
<td>canonical ensemble</td>
</tr>
<tr>
<td>PAW</td>
<td>Projector Augmented Wave</td>
</tr>
<tr>
<td>PBE</td>
<td>Perdew–Burke–Ernzerhof</td>
</tr>
<tr>
<td>PES</td>
<td>Potential Energy Surface</td>
</tr>
<tr>
<td>PFC</td>
<td>Plasma Facing Components</td>
</tr>
<tr>
<td>PFM</td>
<td>Plasma Facing Materials</td>
</tr>
<tr>
<td>PW</td>
<td>Plane Waves</td>
</tr>
<tr>
<td>PWI</td>
<td>Plasma Wall Interaction</td>
</tr>
<tr>
<td>RMSE</td>
<td>Root Mean Square Error</td>
</tr>
<tr>
<td>SCF</td>
<td>Self-Consistent Field</td>
</tr>
<tr>
<td>TDS</td>
<td>Thermal Desorption Spectroscopy</td>
</tr>
<tr>
<td>TFTR</td>
<td>Tokamak Fusion Test Reactor</td>
</tr>
<tr>
<td>VASP</td>
<td>Vienna Ab Initio Simulation Package</td>
</tr>
<tr>
<td>VDE</td>
<td>Vertical Displacement Events</td>
</tr>
<tr>
<td>W</td>
<td>Tungsten</td>
</tr>
<tr>
<td>WEST</td>
<td>W-tungsten Environment Steady-state Tokamak</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>----------------------</td>
</tr>
<tr>
<td>ZPE</td>
<td>Zero-point energy</td>
</tr>
</tbody>
</table>
1 List of publications

The following articles are an integral part of this thesis and they are reprinted in the annex at the end of the thesis.

- Performance of DFT functionals for properties of small molecules containing beryllium, tungsten and hydrogen.
  
  **L. Chen**, D. Süß, I. Sukuba, M. Schauperl, M. Probst, T. Maihom and A. Kaiser
  Nuclear Materials and Energy (2020) 100731

- Beryllium, tungsten and their alloys Be$_2$W and Be$_{12}$W: Surface defect energetics from density functional theory calculations.
  
  **L. Chen**, I. Sukuba, M. Probst and A. Kaiser
  Nuclear Materials and Energy 16 (2018) 149-157
  https://doi.org/10.1016/j.nme.2018.06.021 [2].

- A neural network interface for DL_POLY and its application to liquid water.
  
  I. Sukuba, **L. Chen**, M. Probst and A. Kaiser
  Molecular Simulation (2018) 1-6
  https://doi.org/10.1080/08927022.2018.1560440 [3].

- Iterative training set refinement enables reactive molecular dynamics via machine learned forces
  
  **L. Chen**, I. Sukuba, M. Probst and A. Kaiser
  RSC Advances 10(8) (2020) 4293-4299
The author also contributed to the following publications that are not directly related to the topic of this thesis.

- **Thermal Analysis on Various Design Concepts of ITER Divertor Langmuir Probes.**
  
  **L. Chen**, W. Zhao, G. Zhong, C. Watts, James P. Gunn, X. Liu, Y. Lian & DLP Team
  
  Fusion Science and Technology, 73 (2018) 568-578
  
  https://doi.org/10.1080/15361055.2017.1415614 [5].

- **Highly Stable [C60AuC60]+/− Dumbbells.**
  
  M. Goulart, M. Kuhn, P. Martini, **L. Chen**, F. Hagelberg, A. Kaiser, P. Scheier, and A. M. Ellis
  
  
  https://pubs.acs.org/doi/10.1021/acs.jpclett.8b01047 [6].

- **Clusters of betaine with positive and negative ions: Evidence for the betaine tetramer being magic.**
  
  
  Journal of Chemical Physics, 151(18) (2019) 184303
  
  https://doi.org/10.1063/1.5124553 [7].

The work for publications [1]-[4],[6] was carried out at the University of Innsbruck. The work for article [5] was performed during a research stay at the Southwestern Institute of Physics (SWIP) in China. The work for article [7] was performed partly during a research stay at the Bar Ilan University in Israel.
2 Introduction

In the past several decades, the energy demand in the world has grown rapidly. Despite the emergence of renewable energy, such as solar, wind, tides, geothermal energy, etc., the traditional fossil fuels like coal, oil or gas are still the primary energy source. Burning the fossil fuels drives up CO\textsubscript{2} emission and leads to global warming. The global climate change motivates us to develop safe, environmental friendly and sustainable energy sources. Although it will take decades to achieve commercial power generation, fusion energy can become a promising option to meet all these requirements.

2.1 Nuclear Fusion

Fusion energy is nuclear energy released when light nuclei fuse to bigger products. Table 2.1 gives an estimate of energy released from 1 kg coal, 1 kg U\textsubscript{235} in fission and 1 kg of a deuterium-tritium mixture in a fusion reaction. With the same amount of fuel, D-T fusion reaction can release about a factor of five more energy than the U\textsubscript{235} fission reaction. The energy obtained from fusing D and T and burning coal differ by eight orders of magnitude.

<table>
<thead>
<tr>
<th>Resources</th>
<th>Power (kwh)</th>
</tr>
</thead>
<tbody>
<tr>
<td>coal</td>
<td>10</td>
</tr>
<tr>
<td>U\textsuperscript{235}</td>
<td>23,000,000</td>
</tr>
<tr>
<td>D-T</td>
<td>120,000,000</td>
</tr>
</tbody>
</table>

Table 2.1: Estimates of energy produced from 1kg various resources [8].

Fusion reactions neither produce CO\textsubscript{2} nor any harmful chemicals. The most critical environmental issue in fusion are the high-energy (up to 14 MeV in D-T reaction) neutron products. In ITER, these neutrons are expected to be slowed and captured in the blanket and thus they will not pose threat to the outside environment [9]. However, as they pass
through the structural material on their way to the blanket, the neutrons can activate these materials. However, the unstable isotopes of the structural material have short half-life times and their nuclear waste storage time is about 100 years, much shorter than for fission waste. For example, Pu\(^{239}\) created by transmutation of uranium atoms in fission reactors has a half-life of 24,110 years. Therefore, if we consider the entire environmental impact, fusion is a very attractive option in comparison to fossil and fission energy.

Another major advantage of fusion is safety. In comparison to fission, a fusion reaction does not rely on maintaining chain reactions in a large amount of fuel but relies on a continuous input of fuel. With any perturbation of this fuel supply process, the fusion reaction will terminate immediately. Some promising fusion reactions are listed below [10]:

\[
\begin{align*}
D + D &\rightarrow He^3 + n + 3.27 MeV \quad (2.1) \\
D + D &\rightarrow T + p + 4.03 MeV \quad (2.2) \\
D + T &\rightarrow He^4 + n + 17.6 MeV \quad (2.3) \\
D + He^3 &\rightarrow He^4 + p + 18.3 MeV \quad (2.4)
\end{align*}
\]

where D and T are the hydrogen isotopes deuterium and tritium respectively, n is a neutron and p is a proton. The D-D reaction (2.2) is the most desirable fusion reaction in the sense of unlimited source of fuel, since D can be extracted from Earth seawater. However, it is very difficult to achieve this reaction on earth, because its cross section is very low. Figure 2.1 shows the experimentally measured cross sections of D-D, D-T and D-He\(^3\) reactions (Eqns. (2.2)(2.3)(2.4)).

The cross section of the D-He\(^3\) reaction is a bit higher than for the D-D reaction. The main challenge of this reaction is that He\(^3\) as a reactant does not exist in the natural environment. In spite of that, the D-He\(^3\) reaction is still attractive. Compared to all other reactions, this reaction releases the largest amount of energy. Moreover, it does not have neutron products, hence the problems associated with the materials damage induced by neutron irradiation could be largely reduced, and its main product He\(^4\) has advantages concerning the extraction of energy.

From Figure 2.1, the cross section of the D-T reaction increases dramatically from an energy of 10 eV and achieves a maximum of about 5 barns at the deuteron energy \(K_D\)
\(\approx 120\) keV. However, in the low energy regime the cross sections of D-He\(^3\) and D-D are zero. The combination of the potential energy surface of the two colliding nuclei can lead to a resonance. For the D-T interaction, this resonant process enhances the probability of a fusion reaction occurring, corresponding to a higher cross section. Taking both the amount of energy released and the conditions to initiate the fusion reaction into consideration, the D-T reaction is regarded as the most efficient fusion reaction that can be realized in the laboratory.

\[
\begin{align*}
\text{D-T} & \quad \text{D-D} \\
\text{D-He}^3 & \quad \text{D-He}^3
\end{align*}
\]

**Figure 2.1:** Measured cross sections \(\sigma\) of D-D, D-T and D-He\(^3\) fusion reactions [11]. This figure is reproduced with permission from Krieger Publishing.

D-T reactions require a continuous supply of tritium to sustain the reaction, but T is extremely rare in nature. It is radioactive with a short half-life of 12.3 years [12], and in consequence almost no natural T exists on earth. The global inventory of T produced in nuclear fission reactors with moderation by heavy water is approximately 20 kg per year, while a fusion power plant will require 100\(~\sim\)200 kg T per year.

In order to test the in situ production of tritium in a fusion environment, a so-called test blanket module (TBM) is planned to be installed in ITER. The experimental data will provide an insight into achieving tritium self-sufficiency, a compulsory requirement for a future power reactor [13]. Reactions (2.5) and (2.6) are two ways to breed tritium with the former producing energy and the latter consuming energy.
Although it is regarded to be the easiest fusion reaction to achieve on earth, the D-T fuels still have to be heated up to 150 million K, which is almost ten times hotter than the fusion reaction occurring in the sun. At such a high temperature, the electrons in the atoms are dissociated from the atomic orbitals, resulting in a fluid consisting of free electrons and free nuclei known as plasma. Confining the hot plasma in a finite volume is one of the most important issues of controlled fusion reactions. Since plasma is electronically conductive, it can be confined by electromagnetic fields. The Tokamak and the stellarator are two promising types of devices using magnetic fields to confine the hot plasma. In fusion devices, a twisted magnetic field is required to confine the plasma. One significant difference for the tokamak and the stellarator is the way to generate this twisting magnetic field. The tokamak relies on a toroidal field generated by external coils together with a poloidal field generated by the plasma current, while the stellarator relies on magnetic fields entirely generated by external coils [14]. The latter could be handled satisfactorily only recently and from an engineering point of view, it is much easier to construct a tokamak than a stellarator, while, in terms of plasma stability and plasma confinement, the stellarator is superior. Thus fusion research has been performed in parallel in both tokamaks and stellarators.

Notably, Wendelstein 7-X in Germany, the world’s largest stellarator, has achieved a world record of the fusion products (plasma density, plasma temperature and the pulse time) in the stellarator type of device. Wendelstein 7-X is designed to demonstrate a continuous plasma operation for application in fusion power plants, while no fusion reactions are planned in this device [15]. Due to the high technical complexity, fewer stellarator devices have been constructed compared to tokamaks, and the latter is currently the leading candidate for a demonstration fusion power plant.

2.2 ITER project

An international collaboration has been established to achieve fusion on earth. The Joint European Torus (JET) tokamak was built by the European Union in Culham, United
Kingdom. JET began operation in 1984 and it is still the world’s largest operational magnetic confined fusion device. Remarkable achievements have been made in JET. For instance, JET is currently the only tokamak capable of operating on D-T fuel. Within this D-T operation, JET set the current world record for Q (the ratio of the amount of the power produced compared to the input power) at 0.67 in 1997 [16].

The idea of the ITER was proposed by General Secretary of the Soviet Union Mikhail Gorbachev and the president of United State Ronald Reagan at Geneva in 1985, and it aimed to obtain fusion energy strongly based on a wide international cooperation. A detailed engineering design of ITER was finished in 2001, 16 years after the Geneva meeting. The ITER agreement was finally signed by the seven members (China, EU, Indian, Japan, Korea, Russian and the United States) at Paris in France in 2006, and it was agreed to construct ITER in Saint Paul-lez-Durance, France [17].

The ITER tokamak was designed for several goals, i.e., producing 500 MW energy from a 50 MW energy input (Q=10), achieving a sustained D-T plasma of 400 s, testing tritium breeding and demonstrating the safety issues of a fusion device. The first plasma in ITER is expected in 2025 and D-T operations will begin in 2035 [13].

### 2.3 Plasma facing materials

A cut of the designed ITER vessel is shown in Figure 2.2. The blanket module, which covers most of the inner vacuum vessel (around 620 m²), consists of detachable first wall panels directly facing the plasma and the shield block that is designed for neutron shielding. The divertor is placed at the lower part of the vacuum vessel, which consists of a stainless steel supporting structure and three plasma facing components: the inner and outer vertical targets and the dome. It is designed to extract heat and helium ash produced by the fusion reaction, and to protect the surrounding components from the neutron and heat irradiation.

The Plasma Facing Components (PFC) in ITER will be exposed to steady state heat loads of 2-5 MW/m² at the first wall and of 10-20 MW/m² at the divertor. Apart from the steady state thermal loads, they will also suffer short (0.5 ms), but intense and high heat loads up to GW/m² during ELMs occurring in the H-mode operation. Additionally, off-normal events such as disruptions (with the heat load of up to 30 MJ/m² for 2-5 ms) and VDE (with the heat load of up to 60 MJ/m² for 100-300 ms) can also happen due
2.3 Plasma facing materials

Figure 2.2: Cross section of the designed ITER vacuum vessel [18].

to the loss of plasma control [19]. Therefore, the PFM, armoring the first wall and the divertor, must have a sufficient heat transfer capacity.

Under transient or off-normal loads, the PFM will be eroded and enter the plasma as impurity, hence another key requirement of PFM is good compatibility with plasma. Be will be used as armor material of the first wall in the ITER Tokamak [20] because it has a favorable combination of low-Z (atomic number), high thermal conductivity, good plasma compatibility and high affinity towards reaction with oxygen [21]. However, Be has the disadvantages of a low melting point (1560 K) and a low surface binding energy (3.32 eV [22]), which leads to a higher physical sputtering yield. Due to the toxicity of beryllium-containing dust, beryllium safety is a critical issue in the manufacturing process of first wall panels and in ITER operation. This also limits the usage of Be as plasma facing materials in currently existing fusion devices. JET is the only operational tokamak employing Be as PFM.

The divertor in ITER was initially designed as a component including both Carbon Fibre Composite (CFC) and W armor in the H/He plasma operation phase [23]. CFC have an
excellent thermal conductivity of up to 400 W/(m·K), leading to a high resistance to extreme thermal loads. In addition, carbon is a low Z element. Although the erosion of CFC induced by localized overheating leads to the degradation of the plasma confinement, it will not lead to plasma disruption [24]. These advantages make carbon the most widely used PFM in today’s fusion devices.

However, a major drawback of carbon materials is a high level of fuel retention by co-deposition of hydrocarbon complexes. With the design of a CFC/W divertor, the retained T in the inner wall would quickly reach the allowed limit of 700 g tritium set by the French nuclear safety authorities in the ITER D-T operation phase [25]. Moreover, with the exposure to energetic neutron irradiation, the thermal conductivity of carbon will be significantly decreased. For these two reasons, carbon has been eliminated from the candidates of PFM of ITER and future fusion reactors. The ITER CFC/W divertor design was switched to a full-W design in 2013 [26].

Tungsten has the highest melting point (3695 K) among all elements, combined with a relative high thermal conductivity (approximately 160 W·m⁻¹·K⁻¹) [19], which gives a rise to a superior resistance to the high heat loads in the divertor region. In contrast to carbon, the retention of fusion fuel in tungsten is not a big problem. The JET ITER-Like Wall experiment concluded that the T retention rates in the full metal plasma facing wall consisting of Be and W was reduced by a factor of ten compared to the full carbon wall [27]. Besides, several plasma campaigns in the Axially Symmetric Divertor Experiment (ASDEX) Upgrade equipped with full tungsten PFM have shown a reduction of long time fuel retention by more than a factor of five compared to the previous operations with carbon being the dominant wall material [28].

Compared with the low-Z materials beryllium and carbon, tungsten has a far lower sputtering yield. In order to get a direct comparison of the erosion of carbon and tungsten in a fusion environment, during the 2004-2005 operation campaign, some carbon tiles of the divertor in ASDEX Upgrade were coated with tungsten. The net erosion rate of carbon was found to be about 10-20 times higher than that of tungsten [29]. Nowadays, tungsten is regarded as one of the most promising PFM in future fusion power plants. However, tungsten certainly has some shortcomings. For example, it will transmute into rhenium (Re) under the 14 MeV neutron irradiation [30]. This will change the composition of the material and the new element may degrade the ductility of the tungsten material. Another serious problem is the high radiation heat loss induced by tungsten impurities.
in the core plasma [31].
To investigate the influence of tungsten impurities on plasma operation, several tokamaks have updated their plasma facing materials into tungsten. Since the 2007 campaign, all plasma facing components in ASDEX Upgrade have been equipped with tungsten-coated graphite tiles [32]. The ITER-like wall was installed in the JET tokamak in 2011 [33], with the divertor consisting of bulk tungsten and W-coated CFC tiles without active cooling system and the first wall consisting of Be tiles. In 2014, the upper divertor of Experimental Advanced Superconducting Tokamak (EAST) has been changed into an ITER-like tungsten divertor with actively cooled W/CuCrZr monoblock components for the striking points of both inner and outer vertical targets, and with W/CuCrZr flat-type components for the DOME and the baffle parts of both inner and outer vertical targets [34]. In the frame of the W-tungsten Environment Steady-state Tokamak (WEST) project launched in 2013 [35], the Tore Supra divertor will be armored with tungsten coating or with tungsten monoblocks.

2.4 Plasma wall interaction

The erosion and re-deposition of PFM under high thermal particle loads directly affects the life time of the inner wall components in fusion devices, and conversely eroded wall materials will result in a radiative power loss, a dilution of the plasma and a disturbance of the stability of plasma confinement. Additionally, it is necessary to remain within the safety limit for the T retention in the in-vessel components. All these issues shall be determined by the study of plasma wall interactions. Damages to PFM induced by neutron irradiation are outside the research topic of this thesis.

2.4.1 Erosion

The wall materials in ITER can be sputtered by the impacts of various type of particles, including H isotopes and He as the fusion fuels and products, the noble gases Ne and Ar used for the radiation cooling in the divertor region, oxygen and carbon as plasma impurities and Be being sputtered and then re-deposited on the surface as energetic particles.
Physical sputtering is a well-known erosion mechanism under PWI. It is the kinetic ejection of surface atoms resulting from the momentum transfer from incident particles to target atoms in the collision process. In an elastic collisions process as shown in Figure 2.3, based on energy and momentum conservation the maximum energy transfer in the first collision from the incident particle with the mass $m$ to the target atom with the mass $M$ is $\frac{4mME}{(m+M)^2}$. Here $E$ is the kinetic energy of the incident particle. The energy transfer in the second collision is $E(1 - \frac{4mM}{(m+M)^2})\frac{4mM}{(m+M)^2}$. Surface atoms can be sputtered only if the remaining kinetic energy in the collision is sufficient to overcome the surface binding energy $E_s$. If we assume that at least two collisions are necessary for light particles incident on heavy materials, the sputtering threshold $E_{th}$ can simply be approximated by [21]:

$$E_{th} = E_s \frac{(m + M)^4}{4mM(m - M)^2} \quad (2.7)$$

Estimates of the physical sputtering threshold using Eqn. (2.7) are given in Table 2.2 for the Be and W surface by energetic impacts of H isotopes and for Be impacting on W.

From Table 2.2, the estimated sputtering thresholds for H, D or T impacts on the
W surface are much larger than on the Be surface. Therefore, compared to Be, W is more resistant to physical sputtering by the impacts of ions or neutrals from the plasma. However, an estimated sputtering threshold of 60 eV indicates that W atoms could be sputtered by energetic Be in the divertor region.

The sputtering yield, i.e. the average number of atoms ejected from the target surface per incident particle, depends on the incident angle, the incident energy, the surface temperature, the surface roughness and the incident particle mass. The sputtering yield depending on the incident energy $E_0$ at normal incident can be analytically described by the Bohdansky formula [37].

$$Y(E_0) = qS_{nTF}(\epsilon)[1 - \left(\frac{E_{th}}{E_0}\right)^2](1 - \frac{E_{th}}{E_0})^2$$  \hspace{1cm} (2.8)

where $S_{nTF}(\epsilon)$ is the Thomas-Fermi nuclear stopping cross section, $\epsilon$ is the reduced energy and $q$ is a fitted parameter. The Bohdansky formula is based on Sigmund’s results [38] for backward sputtering in the approximation of no inelastic losses, with two modifications concerning the deposited energy and the momentum distribution of the recoiling target atoms [37].

The angular dependence of sputtering yield can be computed by the Yamamura formula where the sputtering yield at the normal incident $Y(E_0)$ is multiplied by an angular dependent term [39]:

$$Y(E_0, \alpha) = Y(E_0)[cos\alpha]^{-f}exp(f[1 - \frac{1}{cos\alpha}sin(\eta)])$$  \hspace{1cm} (2.9)

With $\alpha$ being the incident angle, $f$ and $\eta$ being the fitted parameters.

The Bohdansky and Yamamura formulas are usually used to fit the experimentally observed or computer simulated sputtering yields under fusion environment. In the 3-D
Monte Carlo code ERO, the physical sputtering yield of the particles (Be or W) sputtered from the plasma facing surface are also calculated based on these two formulas [40].

In 2003, Eckstein and Preuss proposed a new empirical formula (known as Eckstein formula), which describe energy and angular dependencies of the physical sputtering yield [41].

\[ Y(E_0) = qS_n^{K-C}(\epsilon) \frac{\left( \frac{E_0}{E_{th}} - 1 \right)^\mu}{\lambda + \left( \frac{E_0}{E_{th}} - 1 \right)^\mu} \] (2.10)

In the Eckstein formula, the fitting parameter from the Bohdansky formula remains and the stopping cross section \( S_n^{TF}(\epsilon) \) is replaced by \( S_n^{K-C}(\epsilon) \) based on the Kr-C potential [42]. The threshold energy \( E_{th} \) appears in the nominator and in the denominator to ensure that the yield at high energy is not affected by \( E_{th} \). Two additional parameters \( \lambda \) and \( \mu \) are introduced. In comparison with the Bohdansky formula, the Eckstein formula gives a better description of the yield at low incident energies near the threshold [41]. It also gave a better agreement with experimental and simulated data for the case of tungsten self-sputtering than the Bohdansky formula [42]. In Ref. [4], we have used the Eckstein formula to fit the Be self-sputtering yield obtained from computational simulations at low energies and the experimental data at high energies.

**Erosion of beryllium**

In a fusion reactor, the life time of PFC is determined by the net erosion of PFM that is defined as the difference between gross erosion and re-deposition. Gross erosion indicates the amount of PFM sputtered from the surface and entering the plasma as impurities, while re-deposition represents the amount of deposited material from the plasma [29].

The main disadvantage of a low-Z material like Be is its high erosion rate caused by physical sputtering, which largely limits the life time of Be components. Since Be has a very high affinity to oxygen, sputtering experiments on a Be surface need to be carried out in high vacuum. A very early experimental study of Be erosion under H, D and He ion impacts with a wide range of energies (50 eV-8 keV) was performed by Roth et al. in 1979. It was carried out in a high current ion accelerator. The sputtering yield showed a peak near 1 keV for all three impacts, with an amount of 1-3% for H and D and 10% for He impacts, but the yield decreased sharply at lower energies. The sputtering thresholds of 24 eV for H and of 33 eV for D were obtained by fitting an empirical formula [43].
2.4 Plasma wall interaction

To supplement the experimental data, computational simulations were performed by the Monte-Carlo codes TRIM.SP [44] and ACAT [45]. The energy dependence of the sputtering yield in numerical simulations showed a maximum around 200 eV, amounting to 2-4% for H, D and T impacts [44, 45].

Ten years later, another experiment by Roth et al. investigated the sputtering yield of Be by D and heavy ion (He, Ne, Ar) impacts. The influence of surface oxidation, roughness, temperature, and incident angle was discussed. In the case of D impact, the sputtering yield showed a dramatic increase from room temperature to 920 K at energies lower than 1 keV. Above 1 keV, the surface temperature did not have a big influence on the yield. The sputtering yield was found to be larger at grazing angle than at normal angle. Differences between experimental and computational sputtering yields on angle dependence were attributed to the fact that no surface roughness effects were included in the simulations. An extrapolation by TRIM-SP simulations showed that the Be self-sputtering yield could be larger than 100% at energies above 100 eV with incident angles between 45° to 75° [46].

Experimental data of Be self-sputtering by Be ion impacts at high energies (1.5-10 keV) was later obtained by Guseva et al [47]. The self-sputtering yield of Be with a surface temperature of 673 K had a maximum value of 31% at an ion energy of 1.5 keV [47]. Another work by Guseva et al. pointed out that the Be self-sputtering yield at a projectile energy of 0.9 keV does not depend on the target temperature. The studied temperature range was from 370 to 1070 K [48]. MD simulations of Be self-sputtering at low impact energies (50-100 eV) can be found in [49, 50], where an empirical pair potential and a bond-order potential have been employed. The MD simulated data in [49, 50] was in good agreements with previous Monte-Carlo calculations, but no experimental data at this low impact energy range existed for a direct comparison.

Chemical sputtering is the ejection of surface atoms by chemical reactions with incident particles, forming volatile reactants that can desorb from the surface. In fusion devices, carbon surfaces erode chemically via hydrocarbon (CH₃C complexes) emission and its reaction mechanism was introduced in detail by Küppers et al. [51, 52]. The observation of BeD spectra in the divertor region of the JET tokamak in 1998 indicated the possibility of a chemical sputtering from the beryllium surface [53]. Experimental simulations in the linear plasma device PISCES-B showed that the presence of BeD in front of Be targets was a product of chemical sputtering of Be under D impacts. Various electron
temperatures (5 to 30 eV) and ion fluxes (0.8 to 6.0 ×10^{22} \text{ m}^{-2}\text{s}^{-1}) were chosen to simulate edge plasma relevant conditions. The chemical sputtering yield of Be, depending on the surface temperatures, was observed to decrease with increasing ion flux, which is opposite to physical sputtering [54]. This trend was confirmed by MD simulations using a bond-order potential and the mechanism of chemical sputtering was explained. The simulations showed that chemical effects contributed to almost all the total Be sputtering at low ion energies (7-20 eV), and the fraction of chemical sputtering turned out to be smaller at higher ion energy. Likewise, experiments performed in PISCES-B suggested that the fraction decreased from 80% to 40% when increasing the projectile energy from 9 eV to 70 eV [55].

Beryllium sputtering experiments in a tokamak environment were systematically carried out in JET with various limiter configurations, aiming to evaluate the life time of first wall components in ITER. The sputtering yield under D impact with the energy ranging from 25 to 175 eV was determined by in situ emission spectroscopy observations [56]. At low impact energy (25 eV), the gross sputtering yield was lower than 4.5%, and the corresponding net sputtering yield was determined as 2.3%. Based on post-mortem analysis, the net erosion of Be tiles was approximately half of the gross erosion. At a medium impact energy of 75 eV, the gross sputtering yield of about 9% had two contributors: two third resulting from physical sputtering and one third from chemical sputtering that was confirmed by the observation of BeD emission [56]. This agreed well with MD simulations by Björkas et al [55]. A strong reduction of chemical sputtering was measured with increasing the Be surface temperature. For energies larger than 150 eV, the erosion was dominated by Be self-sputtering. The total sputtering yield increased dramatically with impact energy from 150 eV to 200 eV, leading to a yield far larger than 100% [56], which means that more than one atom was ejected from the target surface per incident projectile. In summary, at higher ion energies, Be self-sputtering can be a serious problem that limits the life time of Be PFC.

**Erosion of tungsten**

In contrast to beryllium, tungsten erosion mechanisms have been more intensively investigated in tokamak experiments in recent years. The erosion rate was locally larger than 0.06 nm/s in the strike point region of the outer W divertor during the 2004-2005 charge period in ASDEX Upgrade. It was found that the tungsten coating was completely
eroded from the rough surface, but most of it re-deposited in the recessions and pores of the carbon bases [29].

During the 2016 spring experimental campaign in EAST with the ITER-like tungsten divertor [57], the sputtering yield of tungsten was found to be higher than in JET [58] but close to the one in ASDEX Upgrade [31]. The higher sputtering observed in EAST was mainly attributed to carbon ion bombardment [57]. In EAST, carbon is the armor material of the limiter and lower divertor region. Once carbon is sputtered, it will re-deposit on the tungsten surface with high impact energy. Lithium aerosol injection was found to effectively cool down the edge plasma and suppress W erosion [57]. Tungsten erosion in JET was mainly caused by bombardment with beryllium, the dominant impurity in the plasma [58]. Tungsten erosion and re-deposition have also been studied by exposing W-coated graphite tiles to the divertor plasma discharge in DIII-D, a tokamak operated since 1980s by General Atomics (GA) in the USA. The net erosion rate was found to be 0.26 nm/s for the 1 mm coated W samples. In agreement with predictions from ERO modelings, the net erosion rate could be reduced to be 0.16 nm/s with D\textsubscript{2} gas puffing in the divertor region [59].

The latest experiment in JET with H-mode D plasma operations found that about 94% of the eroded W by ELM loads re-deposited on the divertor region, and the ERO modeling has reproduced the experimental results well. The high re-deposition ensured a low W impurity concentration in the confined plasma [60].

Apart from the tokamak experiments, tungsten erosion caused by high energy transient loads like ELM, VDE and disruption was studied in the plasma gun facility [61] and in electron beam facilities [62, 63]. These experimental findings showed that under ELM transient loads, crack formation was a critical problem of W, due to its high brittleness. For VDE and disruption loads, surface melting was the main factor to cause W damage [61, 62, 63].

Chemical erosion of tungsten was observed by oxygen impacts at an energy of 150 eV in a temperature range of 1200-1900 K via the formation of the WO\textsubscript{3} molecule [64]. At a temperature above 1900 K, sputtering of pure tungsten prevailed.

**Mixed-material effects**

In ITER, eroded Be from the first wall may be transported by the magnetic field, and then deposited on the W divertor surface. Further penetration into the W bulk can lead
to formation of Be-W mixed materials, notably Be$_2$W, Be$_{12}$W and Be$_{22}$W alloys have been observed in experiments in linear plasma or ion beam facilities [65, 66, 67, 68, 69, 70]. As a consequence, the existence of such surface layers may change the physical properties of the PFM, especially decreasing the high melting point of W [65, 66, 67]. A recent test in PISCES-B indicated that the Be-W mixed layer weakened the thermal resistance of W to ELM like transient loads [71]. Major parameters including surface binding energies, adsorption energies of Be$_2$W and Be$_{12}$W govern the physical sputtering and adsorption process and were calculated by DFT in our group [2, 72].

2.4.2 Hydrogen isotopes in PFM

The inventory of tritium in the vacuum vessel and a further permeation into the cooling water is a critical concern for future fusion reactors [73].

The retention of tritium in graphite is a severe problem in fusion devices. The earlier mixed deuterium-tritium plasma campaigns in JET [74] and Tokamak Fusion Test Reactor (TFTR) [75] operating with carbon limiters showed that 30-40% of the tritium input was accumulated in the machine only after a few days. This is the main reason why carbon cannot be used as PFM in the D-T operation in ITER and in the fusion power plant. A maximum D retention rate of $1.5 \times 10^{20}$ D/s was observed in the JET ITER-Like Wall experiments using beryllium and tungsten as PFM. The observed fuel retention was reduced by a factor of 10 in comparison with carbon-wall operations [76]. This experimental finding confirmed the necessity to changing from CFC/W to a full W divertor design in ITER.

Diffusion and retention of H isotopes in Be

H isotopes diffusion in beryllium was studied experimentally since 1967 [77, 78]. The diffusion rate $D$ follows an Arrhenius equation $D = D_0 e^{-E_D/kT}$, where $D_0$ is the pre-exponential factor and $E_D$ is the activation energy. Jones and Gibson reported the pre-exponential factor of $D_0 = 3 \times 10^{-11}$ m$^2$/s and the activation energy of $E_D = 0.19$ eV for T diffusion on Be [77]. In terms of D diffusion on Be, Abramov et al. determined $D_0 = 8 \times 10^{-9}$ m$^2$/s and $E_D = 0.36$ eV [78]. These values are much larger than the ones from Jones and Gibson but the reasons were not discussed in [78].

An early implantation experiment to study H isotopes retention in Be was carried out
by Langley with D having a kinetic energy of 2.5 keV. It reported that almost all D retained in the surface with a flux lower than $2 \times 10^{22}$ D/m$^2$ [79]. Similar experiments were performed by Wampler et al. with D energies between 0.5 and 1.5 keV. It also pointed out a nearly 100% retention for low D exposure (lower than $1.1 \times 10^{17}$ D/m$^2$). For higher D dose, the Be sample was found to be saturated with a relative concentration of 0.31 D/Be [80].

Adsorption of atomic hydrogen isotopes on the Be(0001) surface was first confirmed by Ray et al. using electron energy-loss spectroscopy in 1988 [81, 82]. One year later, R. Yu and K. Lam studied the diffusion of atomic hydrogen by plane-wave basis set DFT calculations on the same Be surface. The most stable adsorption was found at bridge sites, in agreement with the experiments [81]. An interaction energy of 2.39 eV as well as a Be-H bond length of 0.98 Å [83] were derived. Other calculations using the Perdew–Burke–Ernzerhof (PBE) functional with the plane wave basis set yielded a similar binding energy of 2.35 eV [84].

Lossev and Küppers performed a Thermal Desorption Spectroscopy (TDS) experiment to study the adsorption of hydrogen on Be(0001) surfaces in 1993. For the oxygen covered Be surface, molecular H$_2$ desorption was observed at 320 K, which corresponded to an activation energy barrier of 0.61 eV. For a clean surface, the desorption was observed at higher temperature (450 K) with a barrier of 0.87 eV [85]. A plane wave DFT study of H adsorption on the same surface direction in [84] obtained a very similar desorption barrier of 0.90 eV. A recent ab initio study by Bachurin and Vladimirov et al. have studied the effect of H coverage on surface energies of various type of Be surfaces. They reported that the adsorption of H atoms caused a notable reduction of the energy of all studied Be surface [86, 87].

The influence of helium on D retention in Be was studied by exposing Be to D-He mixed plasma experiments [88]. It was pointed out that D retention was increased at low He concentrations (1%), while for high He concentrations (10%) the retention was suppressed [88].

**H isotopes in W**

The problem of H isotopes in tungsten surfaces has been intensively studied by tokamak and linear plasma device experiments together with computational modeling. Nano-size W bubble structures caused by H retention were detected in divertor blankets in JET after
the first ITER-like wall campaign, leading to a degradation of the thermo-mechanical properties of PFM [89]. E. Bernard et al. simulated the synergetic effects of tritium (T) and He radiation on W samples, with T gas deposition after He exposition and demonstrated that increasing the He fluency leads to a higher T inventory. Conversely, lower He flux decreased T trapping [90]. A previous experimental study showed a significant decrease of D retention in W at a temperature of about 573 K by seeding of He in D plasma, where the nano-sized He bubbles were regarded as a diffusion barrier for D [91]. De Backer et al. studied H trapping around defects in W by multiscale modeling, and obtained three regimes of hydrogen solubility, where trapping at dislocation defects only occurred at relatively low temperature [92].

In analogy to beryllium, trapping of atomic H instead of molecular H$_2$ on the W(001) surface was shown by infrared spectroscopy in 1977 [93]. An H diffusion path and a migration barrier of 0.45 eV on a reconstructed W(100) surface were obtained by first-principle calculations [94]. Field ion microscope studies have shown that adsorbed atoms are more mobile on the W(110) plane than on (211) and (321) surface slabs [95]. The dissolution, diffusion, accumulation and bubble formation of hydrogen isotopes in tungsten can directly cause a significant degradation of the thermo-mechanical properties of the material [96].

Ground state electronic properties such as geometries, energetics and vibrational dynamics of small W hydrides that can provide fundamental insight into the reactivity of W and H, have been studied using $ab$ $initio$ calculations [97, 98, 99, 100]. Later neon matrix infrared experiments by Wang et al. compensated the data [101]. Juslin et al. [102] developed a Bond Order Potential (BOP) for the ternary W-C-H system, and the energetics and structures of small W hydrides (WH$_{1-6}$) were reproduced in good agreement with results in Ref. [97, 98, 99, 100, 101].

In EAST experiments, the H inventory in the graphite armor of the lower divertor was found to be significantly larger (beyond a factor of 25) than that in the tungsten armor of the upper divertor [103]. Sang et al. simulated hydrogen isotopes retention in the tungsten divertor of EAST during ELMy H-mode, and they found that the ELMy H-mode operation can increase the amount of the retained hydrogen isotopes by several orders of magnitude due to the ion-induced trap sites and the deeper implantation depth in tungsten [104].

In terms of mixed materials, either experimental or analytical data is very limited. DFT
calculations by Allouche et al. showed that the probability of trapping hydrogen in vacancies in the Be$_{12}$W bulk is higher than in pure beryllium [105].
3 Modeling methods

This section gives an overview of the modeling methods employed in this work. Various DFT functionals with Gaussian type basis set were employed to compare their performance for ground state properties of small molecules containing beryllium, tungsten and hydrogen. The DFT results were compared with couple cluster CCSD(T) results as a reference [1]. The CCSD(T) atomization energies were extrapolated to CBS limits for all studied molecules and all electron calculations were performed for tungsten containing molecules to estimate core electron effects. Plane wave DFT was used to calculate surface defect energies for beryllium, tungsten and their alloys Be$_2$W and Be$_{12}$W [2] and to generate the training data for a NNP [4]. A feedforward artificial Neural Network (NN) was fitted to model energies and forces [3, 4], and self-sputtering simulations on a beryllium surface were performed by both classical MD and ab initio MD [4]. This chapter’s focus is on the theoretical background of MD simulations and the development of potential energy functions using the NN approach. Hartree Fock theory and DFT are reviewed briefly.

3.1 Quantum chemical calculations

3.1.1 Hartree Fock method

Quantum chemical calculations employed in this work are solving the non-relativistic time-independent Schrödinger equation:

\[ H(\vec{r}, \vec{R})\Psi(\vec{r}, \vec{R}) = E\Psi(\vec{r}, \vec{R}) \]  

(3.1)

with various approximations. Electrons move much faster than nuclei because their large mass difference of three orders of magnitude. In the Born-Oppenheimer Approximation [106], electrons are considered to be moving in an environment where the nuclei are fixed. Solving the Schrödinger equation for a molecular system can thus be separated into solving the Schrödinger equations for the nuclei and the electrons (Eqns. (3.2, 3.3, 3.4)).
\[ \Psi(\vec{r}, \vec{R}) = \Psi_{\text{nucl}}(\vec{R})\Psi_{\text{elec}}(\vec{r}, \vec{R}) \quad (3.2) \]

\[ H_{\text{nucl}}\Psi_{\text{nucl}} = E_{\text{nucl}}\Psi_{\text{nucl}} \quad (3.3) \]

\[ H_{\text{elec}}\Psi_{\text{elec}} = E_{\text{elec}}\Psi_{\text{elec}} \quad (3.4) \]

In this approximation, the electron Hamiltonian \( H_{\text{elec}} \) for a system consisting of \( N \) electrons and \( M \) nuclei in atomic units can be written as [107]:

\[ H_{\text{elec}} = -\sum_{i=1}^{N} \frac{1}{2} \nabla_i^2 - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{r_{ij}} \quad (3.5) \]

where \( r_{iA} \) is the distance between the electron and the nucleus, \( r_{ij} \) is the distance between two electrons, and \( Z_A \) is the charge of the nucleus \( A \).

In the Hartree-Fock (HF) theory, the wave-function of the whole system \( \Psi \) can be described as a product of spin-orbital wave-function for each electron [108]. Solving the complicated Schrödinger equation (3.1) for a molecular system is simplified into solving a set of Schrödinger equations for individual electrons.

Electrons are fermions that have to be described by antisymmetric wave-functions. Moreover, the Pauli exclusion principle states that two or more electrons cannot occupy the same spin-orbital [109]. To meet these two requirements, the wave-function can be described as a product of anti-symmetrized single-electron spin orbitals, a Slater determinant [110].

With taking the electron-electron interaction into consideration, the one-electron Hartree Fock equations are derived as [107]:

\[ f(i)\psi_\alpha(\vec{r}_i) = E_i\psi_\alpha(\vec{r}_i) \quad (3.6) \]

\[ f(i) = h(i) + V^{HF}(i) \quad (3.7) \]

\[ V^{HF}(i) = \sum_b J_b(i) - K_b(i) \quad (3.8) \]

where \( f(i) \) is the Fock operator, \( V^{HF}(i) \) is the repulsive potential between the \( i^{th} \) and the other electrons, \( J_b(i) \) is the Coulomb potential, and \( K_b(i) \) is the exchange potential. The Eqns. (3.6)-(3.8) are called the HF equations. They are the basis of many other widely used \textit{ab initio} methods, such as Configuration Interaction (CI), many-body perturbation theory and coupled cluster methods. Within the HF approximation, solving the
many-body problem is substituted by self-consistently solving a series of one electron problems. $V_{HF}(i)$ depends on the wave-functions of all the other $N-1$ electrons, meaning that to solve the Eqn. (3.7) one needs to know the solutions beforehand. Thus the HF equation can only be solved iteratively with the Self-Consistent Field (SCF) approach. In a SCF procedure, an initial guess of spin-orbitals is formulated and used to calculate the Fock-operator $f(i)$ and then to solve the wave function Eqn. (3.7) to obtain a new set of spin-orbitals. With the updated spin-orbitals, the procedure will be repeated until reaching a convergence criterion, namely that $V_{HF}$ and the wave-function change by less than a given threshold.

3.1.2 Complete basis set limit

For very small systems the HF equations may be solved analytically. However, normally it is necessary to expand the spin-orbitals in a basis set. With a complete basis set, the spin-orbital are optimally described without introducing basis set related errors. However a complete basis set usually means that an infinite number of functions must be employed, which cannot be achieved in numerical quantum chemical calculations. In order to check whether the property of interest is converged by increasing the basis sets, an efficient way is to perform the CBS extrapolation. The convergence of the energy with respect to the size of the basis set can be divided into the convergence of HF energy and the convergence of the correlation energy. Helgaker et al. [111, 112] have suggested an empirical formula (3.9) for the CBS limit extrapolation of the correlation energy.

$$E_{\text{corr}}^X = E_{\text{corr}}^\infty + AX^{-3} \quad (3.9)$$

Here $X$ denotes the cardinal number index of the basis set which is 2 for double-zeta, 3 for triple-zeta, etc. $E_{\text{corr}}^X$ denotes the correlation energy calculated using the basis set with cardinal number $X$, $E_{\text{corr}}^\infty$ is the correlation energy at the CBS limit. Although $A$ is an unknown constant, $E_{\text{corr}}^\infty$ can be computed by applying Eqn. (3.9) to a series of calculations with the same type of basis set. Helgaker et al. [111] obtained accurate results for a water system with cc-basis sets. Truhlar proposed the more flexible formula (3.10), introducing an exponent parameter $\beta$. For Ne, HF and the H$_2$O molecule, the optimum
value of $\beta$ for the combination of cc-pVDZ and cc-pVTZ with CCSD or CCSD(T) method was determined to be 2.4 \cite{113}.

$$E^{\infty}_{corr} = E^{\infty}_{corr} + A X^{-\beta}$$

(3.10)

When extrapolating of HF energy, Zhong et al. \cite{114} have proposed formula (3.11):

$$E^{X}_{SCF} = E^{\infty}_{SCF} + B \exp(-\alpha \sqrt{X})$$

(3.11)

Here, $X$ is again the cardinal number, $E^{X}_{SCF}$ is the HF energy calculated using the basis set with cardinal number $X$, $E^{\infty}_{SCF}$ is the HF energy at the CBS limit and $\alpha$ is a constant. The fitted values of the constants $\alpha$ and $\beta$ depend on the basis set and extrapolations and are summarized in \cite{115}.

In Ref. [1], we have performed a CBS limits extrapolation for three types of basis sets def2, cc, and aug-cc using the ORCA code \cite{116}. The atomization energies for small Be/W/H molecules obtained from the CBS limit extrapolation at the CCSD(T) cc-pVQZ/cc-pV5Z level were used as reference energies in the comparison of the performance of 16 DFT functionals.

### 3.1.3 All-electron calculations versus effective core potentials

For elements with a large number of electrons, it is necessary to introduce a model for the core electrons and to treat only the valence electrons explicitly. By employing Effective Core Potentials (ECP) for atoms in the lower part of the periodic table, the computational efforts are substantially reduced, and it often gives good results. Figgen et al. proposed a relativistic ECP to model the $[\text{Kr}]4d^{10}5f^{14}$ cores of the 5d transition metals and they have reproduced all-electron reference data with average deviations of about 0.01 eV \cite{117}. Taking tungsten as an example, in Refs. [1, 2] the 60 core electrons are modelled by Figgen’s ECP \cite{117}, leaving the other 14 valence electrons to be treated explicitly for most calculations.

To estimate the errors from treating core electrons with an ECP, the atomization energies of the tungsten dimer and of tungsten hydride were calculated using an all-electron basis set [1]. According to Einstein’s relativity theory, the mass of a moving electron increases over the rest mass $m_0$:
3.1 Quantum chemical calculations

\[ m = \frac{m_0}{\sqrt{1 - \left(\frac{v}{c}\right)^2}} \] (3.12)

Bohr calculated that the electron in the 1s orbital of hydrogen travels at around \(1/137\) the speed of light \(c\). Then the velocity of a 1s electron in an atom with atomic number \(Z\) can be approximated by the form \(\frac{Z}{137} c\) [118]. For the tungsten atom, the speed of the 1s electron is around 0.54 \(c\), and the electron mass becomes 1.19 \(m_0\). It is clear that relativistic effects cannot be ignored for the core electrons in such a heavy atom. For all electron calculations, segmented all-electron relativistically contracted (SARC) basis sets was employed to replace the ECP for tungsten, and the relativistic effects were considered by using the zeroth-order regular approximation (ZORA) Hamiltonian [119].

3.1.4 Density functional theory

The HF method has limitations because a single slater determinant is insufficient to map the whole Fock space. Additionally, HF theory neglects the correlation energy which is defined as the difference between the exact full CI energy and the HF energy in the limit of a complete basis set.

DFT is among the most popular methods for approximately solving the Schrödinger equation to investigate the electronic structure of molecular and solid state systems. Compared to post-HF methods, the computational cost of DFT is rather low and enables us to approximately study the electronic properties for larger electronic systems including electron correlation energies. Here we only give a brief introduction of DFT, and for a more detailed discussion the reader is encouraged to read the books [120, 121].

DFT is based on the theory that the ground state energy of an electronic system is a functional of the ground state electron density \(\rho(\vec{r})\). The main source of error in DFT calculations comes from the exchange-correlation energy \(E_{xc}\) because its exact form is unknown. Different levels of DFT functionals have been developed to approximate \(E_{xc}\) by using various schemes.

\[ E_{xc} = \int \rho(\vec{r}) \varepsilon_{xc}[\rho(\vec{r})] d\vec{r} \] (3.13)

In Eqn. (3.13), \(E_{xc}\) only depends on the local electron density \(\rho\), thus it is called Local Density Approximation (LDA). The computational low cost LDA is accurate for a system with homogeneous electron gas, while it yields too large atomization energies for...
3.1 Quantum chemical calculations

electronic systems with strongly varying electron density distribution [1]. Improvements over LDA methods must account for the inhomogeneity of the electron density. One step is to add the gradient of the density $\nabla \rho$ to Eqn. (3.13) and is known as the Generalized Gradient Approximation (GGA). Various GGA exchange-correlation functionals have been developed since 1980s. One of the most popular exchange functional B88 proposed by A. D. Becke [122] is constructed by including the contribution of $\nabla \rho$ as a correction to the LDA Slater exchange energy [123]. The B88 exchange functional is usually combined with the correlation functional LYP [124] proposed by Lee, Yang and Parr to produce the well-known BLYP functional. J. P. Perdew et al. also proposed several widely used GGA functionals such as PW86 [125], PW91 [126] and PBE [127]. An extension of GGA methods is to make the exchange-correlation energies depend either on the second order gradient of the density $\nabla^2 \rho$ or on the Kohn-Sham (KS) orbital kinetic energy density $\tau$, which is so-called meta-GGA. For example, with the inclusion of kinetic energy density, the meta-GGA functional $\tau$-HCTH is an extension to the GGA functional HCTH [128].

In hybrid functionals the exchange-correction energies include two contributions, with $E_{xc}$ from GGA or meta-GGA and exact exchange energy from HF. One of the most used hybrid functionals B3LYP [124, 129, 130, 131] is proposed as an improvement over BLYP by addition of HF exchange energy. Hybrid functionals complete the fourth rung of Jacob’s ladder [132], above LDA, GGA and meta-GGA. Double-hybrid methods where the virtual KS orbitals are utilized are considered to complete the fifth. Taking the double hybrid functional B2-LYP proposed by S. Grimme as an example, the exchange energies consist of the GGA (B88) and HF exchange, and the correlation energies consist of the GGA and the perturbative second-order (PT2) correlation [133]. B2-LYP was found to show very promising results for the bond length and harmonic vibrational frequencies for transitional-metal complexes [133].

In common GGA or meta-GGA including hybrids, long-range electron correlations are normally not considered. This leads to a general shortcoming of these functionals for describing van der Waals interactions [134, 135, 136] that play a crucial role in controlling the structure of DNA and proteins, in surface science or molecular films, etc [137]. The damped dispersion correction with the form of $C_6 \cdot f_{\text{damp}}/R^6$ proposed by S. Grimme aims to describe the induced-dipole forces at large interatomic distances [137]. This empirical dispersion energy is added to the total energy. $f_{\text{damp}}$ is a damping function used to avoid
3.1 Quantum chemical calculations

wrong behavior at short distances. We have employed the hybrid GGA functional B3LYP including Grimme D3 dispersion corrections [138] for calculating electronic properties of charged and neutral betaine clusters (Ref. [7]).

Another way to include the long-range correction in the DFT functionals is to separate the exchange energies into short and long-range interactions based on a splitting operator \( \frac{1}{r_{12}} \) [139, 140]. In the range-separated functional CAM-B3LYP, the exchange energy is a combination of 0.19 HF plus 0.81 B88 at short range, and 0.65 HF plus 0.35 B88 at long range [139]. DFT exchange plays a major role at short range, while HF exchange is more important at long range. For the range-separated hybrid meta-GGA functional M11, the percentage of HF exchange at long-range is set to 100% [140].

\( \omega \)B97XD has included both long-range corrections and empirical dispersion corrections, and is considered as a re-optimization of the range-separated functional \( \omega \)B97X by the addition of Grimme D3 dispersion corrections. With respect to \( \omega \)B97X, \( \omega \)B97XD is superior for non-bonded interactions [141].

In Ref. [1] 16 DFT functionals chosen from rungs 1 to 4 of Jacob’s ladder covering long-range and dispersion corrections were employed to compare the performance for properties of small neutral Be/H/W molecules. DFT calculations with Gaussian basis sets, HF and CCSD(T) calculations were carried out in the code Gaussian 16 [142]. CBS limit extrapolations and all electron calculations were carried out in the ORCA programme [116]. The calculations were performed on the high performance compute clusters LEO at the university of Innsbruck.

3.1.5 Density functional theory with plane wave basis sets

The electronic spin-orbitals in plane wave DFT are expanded in a basis set of Plane Waves (PW). The wave-function can be written as a complex function:

\[
\Psi(\vec{r}) = e^{i\vec{k}\vec{r}}
\]

(3.14)

where \( \vec{k} \) is the wave vector. The kinetic energy of a plane wave is \( \frac{k^2 \hbar^2}{2m} \). In order to perform feasible simulations, we are obliged to restrict our basis to a finite set of plane waves, hence plane waves having a kinetic energy above the cut-off value are not considered. To check the convergence of the basis set, a series of calculations with increasing cut-off energy has to be carried out until convergence to a predefined tolerance is reached.
3.2 Classical molecular dynamic simulations

With periodic boundary conditions, the plane wave basis is well suited to describe an infinite system. Therefore, plane wave DFT is normally used for periodic and extended systems like metal bulks or surfaces. When it comes to small molecules, one still needs to construct a sufficient large super cell to avoid interactions of the molecule with its own images or use a modification of the original method. In the case of plane wave DFT calculation for tungsten containing molecules in this thesis, the cut-off energy was set to be 550 eV, approximately corresponding to 55,000 plane waves [121]. The PW basis set is significantly larger than the quadruple-zeta Gaussian type basis set cc-pVQZ [1]. For such cases, it would be more efficient to employ more localized Gaussian or Slater basis sets. On the other hand, plane wave basis sets allow a very efficient evaluation of two-electron integrals and they are free of the Basis Set Superposition Error.

We have used the PBE exchange-correlation function with a PW basis to calculate the surface defect energies for Be, W and the alloys Be\(_2\)W and Be\(_{12}\)W [2] and to generate the training data base for constructing a NNP for a beryllium surface [4]. PBE is a GGA functional from J.P. Perdew’s group [127]. The convergence of the cut-off energy of 550 eV for the same materials was checked in a previous work in our group [72]. All PW calculations were carried out using the Vienna Ab Initio Simulation Package (VASP) code [143, 144]. Most of them were performed on the high performance Vienna Scientific Cluster (VSC-3) and on the LEO clusters in Innsbruck.

### 3.2 Classical molecular dynamic simulations

#### 3.2.1 Newton’s equation of motion

MD simulations are a powerful tool for studying the physical movement of a many-body system at atomistic level. It is commonly used in material surface science, drug design, describing the properties of protein behavior, etc. The atomic trajectories in MD simulation are determined by solving Newton’s equations of motion. Newton’s second law of motion for atom \(i\) for an \(N\)-body system is given as:

\[
\vec{f}_i = m_i \frac{d^2 \vec{r}_i}{dt^2} = -\frac{\partial U(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N)}{\partial \vec{r}_i} \quad (3.15)
\]

\[
\frac{d\vec{q}}{dt} = \frac{\partial H}{\partial \vec{P}} = \frac{\vec{P}}{m} \quad (3.16)
\]
3.2 Classical molecular dynamic simulations

\( m_i \) denotes the mass, \( \vec{r}_i \) denotes the Cartesian coordinates and \( \vec{f}_i \) denotes the forces of atom \( i \). \( U \) represents the potential energy functions. Eqn. (3.16) is the Hamiltonian form of the classical equations of motion. \( H \) is the Hamiltonian corresponding to the total energy of the system, \( \vec{P} \) is the momentum and \( \vec{q} \) is the position vector.

Eqn. (3.15) is a second-order differential equation that can be converted to a first-order differential equation by replacing \( d\vec{r}_i/dt \) with atomic velocities \( \vec{v}_i \) or by replacing \( m_i \) \( d\vec{r}_i/dt \) with atomic momentum. By integrating this first-order equation, the atomic positions \( d\vec{r}_i \) varying with time, i.e. trajectories are determined.

3.2.2 Integrators for the equations of motion

Several numerical methods, such as Verlet-Stoermer integration [145], Runge-Kutta integration [146, 147], Beeman’s algorithm [148], etc., exist for integrating Newton’s equations of motion. Verlet integration has the advantages of good numerical stability and time reversibility. The motion of a particle at time-forward \( \vec{r}(t + h) \) and at time-backward \( \vec{r}(t - h) \) expanded in Taylor series is:

\[
\vec{r}(t + h) = \vec{r}(t) + \vec{v}(t)h + \frac{1}{2} \ddot{\vec{a}}(t)h^2 + \frac{1}{6} \dddot{\vec{r}}(t)h^3 + O(h^4) \tag{3.17}
\]

\[
\vec{r}(t - h) = \vec{r}(t) - \vec{v}(t)h + \frac{1}{2} \ddot{\vec{a}}(t)h^2 - \frac{1}{6} \dddot{\vec{r}}(t)h^3 + O(h^4) \tag{3.18}
\]

where \( h \) is the time increment and \( \dddot{\vec{r}}(t) \) is the third derivative of the position vector with respect to time. By summing two sides of Eqns. (3.17) and (3.18), one arrives at the Verlet integrator:

\[
\vec{r}(t + h) = 2\vec{r}(t) + \ddot{\vec{a}}(t)h^2 - \vec{r}(t - h) + O(h^4) \tag{3.19}
\]

With this integration, the position of an object at time \( t + h \) can be calculated from two previous position vectors at time \( t \) and \( t - h \) respectively and acceleration at time \( t \). The velocities at time \( t \) can be easily computed by subtracting (3.18) from (3.17):

\[
\vec{v}(t) = \frac{\vec{r}(t + h) - \vec{r}(t - h)}{2h} + O(h^3) \tag{3.20}
\]

From Eqns. (3.19) and (3.20), the error in position at time \( t \) of the Verlet algorithm is of the order \( O(h^4) \), and the error in velocity is of the order \( O(h^3) \).
One drawback of the Verlet algorithm is that the velocity cannot be calculated at the same time step as the position because the value of $\vec{v}(t)$ is obtained after $\vec{r}(t + h)$ is calculated. To solve this problem, Swope et al. proposed the velocity form known as Velocity-Verlet algorithm \[149\], which is commonly used to integrate Newton’s equation of motion.

\[
\begin{align*}
\vec{r}(t) &= \vec{r}(t - h) + \vec{v}(t - h)h + \frac{1}{2}\vec{a}(t - h)h^2 \\
\vec{v}(t) &= \vec{v}(t - h) + \frac{h}{2}[\vec{a}(t) + \vec{a}(t - h)]
\end{align*}
\] (3.21) (3.22)

In the Velocity-Verlet algorithm, the velocity appears directly in Eqn. (3.21). The position and velocity at time $t$ can be calculated directly from the position and velocity at time $t - h$. The Velocity-Verlet algorithm as implemented in LAMMPS and VASP was chosen as the time integrator for all MD simulations in Ref. \[4\].

### 3.2.3 Thermodynamic Ensembles

In the micro-canonical ensemble (NVE), the thermodynamic state of the system is characterized by a fixed number of particles, a fixed volume and total energy (kinetic plus potential energy). If the total energy has a significant fluctuation, it indicates that the integration time step is too large. In Ref. \[4\], the time step for the Be self-sputtering simulations was chosen to be 0.2 fs. However, with the same time step, a large energy fluctuation was observed for MD simulations of the Be surface sputtered by D energetic impacts. The reason is that the mass of D is much smaller than the one of Be and therefore at the same kinetic energy of 100 eV, the velocity of D is more than a factor of 2 larger than the velocity of Be. Therefore, a smaller integration step of 0.1 fs was used for D sputtering simulation.

From Eqn. (3.22), the velocity at time $t$ is computed from the velocity at the previous step, and then the instantaneous temperature of the system can be computed from the atomic velocities:

\[
T = \frac{\sum_{i=1}^{N} m_i v_i^2(t)}{k_B f}
\] (3.23)
3.2 Classical molecular dynamic simulations

conserved volume and temperature. To maintain a realistic constant temperature $T_0$, an external heat bath is needed. Consequently, Newton’s equations of motion are modified. In the Nosé–Hoover thermostat [150, 151], the equation (3.15) is modified by a time dependent friction coefficient $\chi$:

$$\frac{d\vec{v}}{dt} = \frac{\vec{f}(t)}{m} - \chi(t)\vec{v}(t)$$

(3.24)

The coefficient is controlled by the instantaneous temperature $T$, and the requested temperature $T_0$:

$$\frac{d\chi(t)}{dt} = \frac{1}{\tau^2} (T - T_0) - 1$$

(3.25)

where $\tau$ is a time constant that is recommended to lie in the range of $0.5 \sim 2$ ps [152]. In Eqn. (3.25), if the instantaneous temperature $T$ is larger than $T_0$, the first derivative of $\chi$ with respect to time become positive. $\chi$ will increase at the next integration step to make the velocity change smaller by Eqn. (3.24) and to decrease the average temperature of the system by Eqn. (3.23).

In Ref. [4], the crystal structure was first equilibrated for 2 ps at 300 K within the NVT ensemble using the Nosé–Hoover thermostat [150, 151]. Consequently, the sputtering simulations were performed within the NVE ensemble.

3.2.4 Force fields

The term force field refers to a combination of parameters and functions describing analytically the interactions in a molecular system. Conservative forces are simply the negative gradient of the potential energy $\vec{F} = -\nabla U$.

Pair potentials describe the simplest interaction occurring between a pair of neighboring atoms, which captures the two main features of atomic forces: repulsive interaction at close range and attractive interaction at larger range. A widely known pair potential is Lennard-Jones potential [153]. The potential energy $U$ for a pair of atoms is given as:

$$U = 4\varepsilon\left[\left(\frac{\sigma}{r}\right)^{12} - \left(\frac{\sigma}{r}\right)^6\right]$$

(3.26)
where $r$ is the interatomic distance, the parameter $\varepsilon$ is the depth of the potential governing the strength of interaction and $\sigma$ is the distance at which the Lennard-Jones potential is zero. The repulsive term in Eqn. (3.26) approximates Pauli repulsion arising from the overlap of electronic orbitals, and the attractive term represents the attraction caused by electron correlations like van der Waals forces. The repulsive term dominates at short distances, while the attractive term dominates at the long distances.

The Lennard-Jones potential proved to accurately model weak van der Waals forces in noble gases [153]. However, its $r^{-12}$ term is without real physical meaning and turns out to be too steep leading to an overestimation of pressure for dense systems. Moreover, the Lennard-Jones potential has only two parameters, which limits its accuracy for fitting the properties of real materials.

To improve the Lennard-Jones potential, the $r^{-12}$ term is sometimes replaced by a weaker repulsive function, like $r^{-9}$. Another modification is the Buckingham potential that replaces the $r^{-12}$ term by an exponential function $\propto e^{-r}$ [154].

The Lennard-Jones potential repels at short distance, attracts at intermediate distances and approaches zero rapidly at large distances. Therefore, the Lennard-Jones potential is a short-range potential. The Coulomb potential in contrast is a long ranged interaction. Ideally all the atoms in the system must be considered to treat long range forces, while a cut-off function can be employed to ignore the atoms at large distance but without losing too much of accuracy. In this work, we have applied periodic boundary conditions on the supercells of solid Be, W and Be/W alloys surface in both DFT and MD simulations [2, 4].

For metals, an improvement over simple pair potentials is a BOP. The form of this potential is motivated by the idea that the bond order depends on the local atomic environment [155, 156]. The interaction energy between atom $i$ and $j$ in Tersoff potentials can be written as:

$$ U_{ij} = f_c(r_{ij})[U_{\text{repulsive}}(r_{ij}) + B \ast U_{\text{attractive}}(r_{ij})] $$

where $f_c$ is a cut-off function, $r_{ij}$ is the distance between atom $i$ and $j$. The second term in Eqn. (3.27) represents the bond. $B$ does not only include the contribution of atoms $i$, $j$ but also the one of other atoms within the sphere of influence of the central atom $i$ determined by cut-off radius.

In BOPs, the coefficient $B$ depends on the local environment of the central atom $i$ and
is introduced to include the effect of variable bond order. In this way the structural chemistry is included in the classical empirical potential [155]. A good transferability of BOPs to systems that have not been used to fit the potential was pointed out in Refs. [155, 156].

3.2.5 Ab initio molecular dynamic simulations

In ab initio MD simulations, the inter-atomic forces are not computed from a high-dimensional potential energy function but computed on-the-fly from electronic structure calculations at every integration step by solving Schrödinger’s equation, while the trajectories are obtained by solving classical Newton’s equations. Ab initio MD does not rely on any fitted parameters, which is accurate but computationally demanding. Nowadays, ab initio MD is still limited to small systems and short time scales.

Born-Oppenheimer Molecular Dynamics (BOMD) is one of the most common types of ab initio MD methods. In BOMD the total energy of an electronic system is minimized with respect to the wave-function \( \psi_i(\vec{r}) \), and then Newton’s equations of motion is given as [157]:

\[
M_I \ddot{\vec{R}}_I = -\nabla_{\vec{R}_I} [\min_{\psi_i(\vec{r})} E(\psi_i(\vec{r}); \vec{R})]
\]  

(3.28)

Calculating the potential energy \( E \) in Eqn. (3.28) involves solving self-consistent equations with a high computational cost. In 1985, Roberto Car and Michele Parrinello proposed the Car-Parrinello method [158] as an alternative to BOMD. In the Car-Parrinello method, a dynamic simulated annealing technique [158] is introduced to minimize the KS functional, and thus a costly self-consistent electron’s wave function minimization is avoided. The Car-Parrinello method is more efficient for large systems. All ab initio MD simulations reported in Ref. [4] are of BOMD type. We have performed BOMD sputtering simulations on a small beryllium surface system (\(~100\) atoms) with a very short integration time (less than 150 fs) to generate the initial training set for constructing a NNP. In order to make a direct comparison with the sputtering yields simulated by NNP, 500 BOMD trajectories with very high computational cost were calculated on a large system (around 500 atoms). All BOMD simulations were carried out using the VASP code.
3.3 Neural network approach

The term Machine Learning (ML) was firstly coined by Arthur Samuel [159]: “it is a field study that gives computers ability to learn without being explicitly programmed”. Machine learning has been widely used in a range of fields such as quantum computing [160], biology [161], medicine [162], data science and business [163]. Figure 3.1 shows the number of peer-reviewed publications using the artificial neural network technique since the 1980s. Applications of machine learning in scientific research increased dramatically during the last 20 years. There are in total around 10,000 publications in the field of chemistry, physics and material science by Nov. 2019.

![Figure 3.1: Number of publications on artificial neural networks from the year 1985 to 2019. The data was extracted from Web of Science [164] in Nov. 2019.](image)

ML has been applied in nuclear fusion research since the 1990s. In 2019, J. Kates-Harbeck et al. [165] proposed a method based on deep learning to predict disruptions. Their approach showed reliable predictions for fusion devices, the data from which has not been used in the training process. The capability of prediction of long warning times
Neural network approach

(3.3) opens up the possibility of active fusion reactor control [165].
ML in representing Potential Energy Surface (PES) has been employed since 1995 [166]. The accuracy of potential functions is directly connected to the reliability of MD simulations. Numerous efforts have been made by generations of scientists to develop potential functions for various systems. Despite of their successful applications, the current potentials still have some shortcomings for describing chemical reactions, arbitrary atomic configurations, dispersion interactions via covalent bonds, etc. [167]
Therefore, developing potential functions is always attractive out of necessity but remains challenging. Using ML approaches to develop an atomistic ‘mathematic potential’ has attracted more and more interest in recent years.
Contrary to traditional empirical potentials based on simple forms to approximate key physical properties, ML potentials do not have a direct physical meaning, but rely on very flexible mathematic forms with many parameters. A ML potential is a multi-dimensional function depending on the atomic positions and on element types. Using these potentials to calculate energies and forces is orders of magnitude faster than using quantum chemistry methods.
ML potentials are fitted by learning energies and forces from sampled “training data”. Normally training data is generated from costly \textit{ab initio} simulations or experimental data. Different ML methods have been proposed to construct potentials, like Gaussian regressions [168, 169, 170], Taylor expansions [171, 172], artificial neural networks [173, 174, 175]. A comparison of these different approaches can be found in [176]. Here, only the feedforward neural network approach is discussed in detail. In this work we have developed NNPs that have been implemented in the sputtering dynamic simulations on the beryllium surface [4].
Artificial neural networks are among the most used ML method. They were originally designed to understand the “all-or-none” character of nervous activity in the human brain in 1943 [177]. With some key developments of introducing hidden layers, activation functions, deep learning, and advanced training algorithms, the NN has turned to be a crucial tool for big data and data analysis in business and internet companies. A NN with only one hidden layer and an arbitrary continuous sigmoidal activation function was proved to be able to approximate arbitrary multidimensional functions in 1989 [178]. In 1995 Blank et al. [166] firstly proposed to use feed-forward NNs with several degrees of freedoms to model the PES, and the NN-based PES has predicted accurate potential
energy for Ni(111) and Si(111) systems. A schematic of feed-forward NN is shown in Figure 3.2, consisting of artificial neurons in the input, the hidden and the output layers. In this design all neurons or nodes in different layers are connected by the weights and the bias parameters. The activation function $f$ is applied to the output of the nodes in the hidden and output layers (see Eqn. (3.29)). Specifically the node $y^j_k$ in layer $j$ is based on the weighted sum of the nodes $y^{j-1}_l$ in the previous layer $j-1$. Here $\omega^{j-1,j}_{l,k}$ is the weight parameter connecting node $l$ in layer $j-1$ to node $k$ in layer $j$, and $f^j_k$ is the activation function that is applied to the outputs in layer $j$ from the inputs in layer $j-1$. Nodes in the same layer are not connected.

$$y^j_k = f^j_k(b^j_k + \sum_{l=1}^{N_{j-1}} \omega^{j-1,j}_{l,k} \cdot y^{j-1}_l)$$  \hspace{1cm} (3.29)$$

Thus the output of a feed-forward NN is an analytical function of the input with an amount of parameters that depends on the architecture of the NN. The parameter space defining the output function can be enlarged by increasing the number of the nodes in the input and in the hidden layers or increasing the number of hidden layers itself.

![Figure 3.2: A schematic of a feed-forward neural network [173].](image)

Instead of a binary activation function (either “on” or “off”) which is present in biological neural networks, continuous activation functions are advantageous for constructing
3.3 Neural network approach

a PES. To implement NNP for MD simulations, the activation function must be differentiable in order to be able to obtain forces. Several activation functions like SoftPlus, Sigmoid, Gaussian, and hyperbolic tangent can be employed.

3.3.1 Behler-Parrinello high dimensional feed forward neural network potential

The total potential energy must not change if the system is translated or rotated and if nuclei of the same elements in the system are exchanged. For representing the PES, the NNP must be invariant to these transformations. This key requirement apparently cannot be satisfied by using atomic coordinates directly as the input of the NN. The NN effectiveness also depends on the size of the input layer and consequently the usage of only interatomic distances and angles will result in a poor interpolation. These problems can be overcome by employing an atom-centered element-specific NN for each atom in the training set. The output of the NN is the energy of atom $E_i$. The total energy $E$ of one configuration can be obtained by summing over all the atomic energies. This idea was proposed by Behler and Parrinello in 2007 [173]. In this approach, the atomic coordinates are converted to symmetry functions [179] describing the local chemical environment. These are used as input nodes in the NN. A cut-off function $f_c$ is introduced to ensure that only the neighbors inside a sphere with the cut-off radius $R_c$ contribute to the atomic energy.

$$f_c(R_{ij}) = \begin{cases} 
0.5[\cos(\pi R_{ij} / R_c) + 1], & R_{ij} \leq R_c \\
0, & R_{ij} > R_c 
\end{cases} \quad (3.30)$$

The cut-off radius $R_c$ should be sufficient to include all energetically relevant neighbor atoms of the central atom. A cut-off radius of 7 Å was used in Ref. [4]. The symmetry function contains the cut-off function to ensure a smooth decay to zero at the cut-off radius.

Behler proposed several types of radial and angular symmetry functions [179], two of which are given in Eqns. (3.31) and (3.32). Radial symmetry functions account for two-body interactions while angular functions account for three-body interactions.
3.3 Neural network approach

\[ G_{i}^{\text{rad}} = \sum_{j \neq i} e^{-\eta (R_{ij} - R_{x})^2} f_{c}(R_{ij}) \] (3.31)

\[ G_{i}^{\text{ang}} = 2^{1-\zeta} \sum_{j,k \neq i,j<k} (1 + \lambda \cos \theta_{ijk}) \zeta e^{-\eta (R_{ij}^2 + R_{ik}^2 + R_{jk}^2)} f_{c}(R_{ij}) f_{c}(R_{ik}) f_{c}(R_{jk}) \] (3.32)

Singraber et al. proposed angular wide symmetry function as given in Eqn. (3.33) [175], which is based on Behler’s angular symmetry function but the factor \( e^{-\eta R_{jk}^2} f_{c}(R_{jk}) \) has been omitted. This function is designed for large angles \( \theta_{ijk} \), where neighbor atoms are far from each atom but close to the central atom.

\[ G_{i}^{\text{ang.w.}} = 2^{1-\zeta} \sum_{j,k \neq i,j<k} (1 + \lambda \cos \theta_{ijk}) \zeta e^{-\eta (R_{ij}^2 + R_{ik}^2)} f_{c}(R_{ij}) f_{c}(R_{ik}) f_{c}(R_{ik}) \] (3.33)

One can see that the symmetry functions depend on the interatomic distances or (and) on the angles. Employing these functions as the input of atomic NN enables representation of a high-dimensional rotationally, translationally and exchange invariant PES. A schematic diagram of a high-dimensional NNP is shown in Figure 3.3.

![Figure 3.3: Schematic of a Behler-Parillo high-dimension feed forward NNP [179].](image)

In the fitting process, the weights and bias are optimized by minimizing the Root Mean Square Error (RMSE) of the energy and forces. The reference atomic forces can directly be computed in \textit{ab initio} methods. For plane wave DFT calculations in the VASP code,
exact Hellmann-Feynman forces acting on atoms are calculated at each ionic step. Due to the motion-invariant plane wave basis, they are the nuclear forces. However, the energy of the atom in its environment cannot be directly obtained from ab initio calculations, thus the total energy of one configuration is utilized as the reference energy. In Behler and Parrinello’s approach, the forces predicted from a NN can be analytically calculated as the negative gradient of the total energy at the atomic positions (Eqn. (3.34)) [180].

\[
F_{ai,x} = -\frac{\partial E_a}{\partial x_{ai}} = -\sum_{\mu=1}^{N_G} \sum_{\nu=1}^{N_a} \frac{\partial E_{\mu}}{\partial G_{\mu\nu}} \cdot \frac{\partial G_{\mu\nu}}{\partial x_{ai}}
\]  

(3.34)

Here, \(F_{ai,x}\) denotes the force of atom \(i\) in configuration \(a\) at \(x\) direction. Two indices \(N_G\) and \(N_a\) represents the number of symmetry functions used in the inputs of atomic NN and the number of atoms in configuration \(a\), respectively.

Since the forces act on each atom and in three directions, the amount of reference forces is much larger than the amount of reference energies in the training set. Normally not all the reference forces are used simultaneously in the actual training process. For example, all the energies but only 2% forces were updated at every epoch to optimize the weights in Ref. [4].

### 3.3.2 Iterative training set refinement

One well-known challenge of NNP is its transferability to structures that are very different from the configurations included in the training set. After generating an initial training set by using ab initio methods, a further iterative refinement may be needed to complete a well-balanced and complete training set.

Artrith and Behler have described a refinement procedure in [181, 180]: To start with, several NNPs are constructed with the same training set; Then one of the fitted NNPs is employed to perform MD or Monte Carlo simulations to generate many new configurations; Next, the forces and energies of these configurations predicted by different NNPs are compared; In the end, the configurations with very different energies or forces from different NNPs are added back to the training set after running a DFT calculation on them. We have implemented this approach to refine the training set (Figure 3.4) for constructing a NNP applied to beryllium self-sputtering simulations [4].
3.3 Neural network approach

Figure 3.4: Schematic of the iterative refinement procedure based on previous work [181, 180] for training set generation. Energies and forces predicted with NNP1 and NNP2 are compared with each other along molecular dynamics trajectories generated with NNP1. The decision (NNP1 ~ NNP2?) of including a particular phase-space point into the training set is made either by the energy criterion |\( E(\text{NNP}_1) - E(\text{NNP}_2) | > \varepsilon_E \) or the force criterion |\( F_{x,y,z}(\text{NNP}_1) - F_{x,y,z}(\text{NNP}_2) | > \varepsilon_F \) [4].

3.3.3 Developing neural network potentials with the n2p2 library

NN library aenet and n2p2

In Ref. [3], we have constructed the NNP for liquid water using the Atomic Energy Network (aenet) package developed by Nongnuch and Urban [182, 183]. Aenet is an open source library that has been successfully applied to accurately predict lattice parameters, energies and bulk moduli for TiO\(_2\) [182], to simulate the complex Li diffusion in crystalline and silicon nanoparticles [184], to construct first-principles phase diagrams for amorphous Li/Si alloy [185], etc. However, we have encountered limitations of using aenet 2.0.3 to construct a NNP for sputtering simulations as described below. Thus in Ref. [4] we
3.3 Neural network approach

worked with another neural network library \textit{n2p2} recently developed by A. Singraber et al. [175, 186]. A comparison of the main settings of these two NN package are given in Table 3.1.

<table>
<thead>
<tr>
<th>Settings</th>
<th>aeonet 2.0.3</th>
<th>\textit{n2p2}</th>
</tr>
</thead>
<tbody>
<tr>
<td>Programming language</td>
<td>Fortran and C</td>
<td>C++</td>
</tr>
<tr>
<td>Training methods</td>
<td>gradient descent method;</td>
<td>gradient descent method;</td>
</tr>
<tr>
<td></td>
<td>Levenberg-Marquardt algorithm [187, 188];</td>
<td>Levenberg-Marquardt algorithm [187, 188];</td>
</tr>
<tr>
<td></td>
<td>Kalman filter method [186]</td>
<td></td>
</tr>
<tr>
<td>Local atomic environment description</td>
<td>Behler-Parrinello symmetry functions [173];</td>
<td>Behler-Parrinello symmetry functions [173];</td>
</tr>
<tr>
<td></td>
<td>smooth overlap of atomic positions (SOAP) descriptor [168];</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Chebyshev descriptor [183]</td>
<td></td>
</tr>
<tr>
<td>Reference data for optimizing the weights</td>
<td>The total potential energy of reference structures</td>
<td>The total potential energy of reference structures and atomic forces</td>
</tr>
</tbody>
</table>

Table 3.1: Comparison of main settings of aeonet 2.0.3 and the \textit{n2p2} package in Dec. 2019.

Aenet includes three different descriptors to describe the local atomic environment, while \textit{n2p2} so far only includes Behler-Parrinello type symmetry functions. In terms of training methods, apart from gradient descent and the Levenberg-Marquardt algorithm [187, 188], \textit{n2p2} also supports a Kalman filter method [186]. An asset of \textit{n2p2} for constructing the NNP applied in MD simulations is that both the total energy of reference structures and atomic forces are used for optimizing the weights in the fitting process.

For a direct comparison, we have used the same training set to construct a NNP for a tungsten surface. The quality of the NNP constructed by aeonet 2.0.3 in which only energy is used for optimizing the weights is compared with the one constructed by \textit{n2p2} in which both energy and forces are used.

The training data comprised 50,000 configurations randomly extracted from BOMD trajectories for 250 fs on a small W(001) surface consisting of 54 atoms. The temperature from 300 to 2000 K varying in different trajectories was controlled by the Nosé-Hoover
3.3 Neural network approach

algorithm [150, 151] with an integration time step of 0.5 fs. Around 90% of the training data was randomly chosen to form the training set to fit the NNP and the rest formed the test set to validate the potential and prevent overfitting.

The iterative training set refinement illustrated in Figure 3.4 was employed to compare the quality of NNPs. NNP1 and NNP2 were fitted using aenet 2.0.3 in which all the energies were updated at each epoch to optimize the weights. The RMSE of the training and test set in the training process with 1000 epochs was monitored and plotted in Figure 3.5. One could see that the errors decrease rapidly at the first 50 steps and then turned to be stable after 500 steps, and overfitting did not happen. The final RMSE in the training and test set was 5.6 and 5.8 meV/atom respectively.

![Figure 3.5: The RMSE of the training and the test set via the training epochs.](image)

NNP3 and NNP4 were fitted using n2p2 in which all the energies and 5% of the forces were updated at each epoch. NNP1 and NNP3 were used to perform a thermostatted MD trajectory to generate some new configurations. The energies of randomly chosen structures along the MD trajectories were recalculated by DFT.

Total energies obtained from two NNPs predictions and from DFT were compared in Figure 3.6. As shown in Figure 3.6 (a), energies predicted by NNP1 and NNP2 are in good agreement between them but significantly differ from the one obtained by DFT for most configurations. This result questions the capability of the developed NNP to correctly represent the physical properties of the studied system. From Figure 3.6 (b), one can find that the energies predicted from NNP3 and NNP4 are in very good agreement with the energies recalculated by DFT, making the NNP considered to be reliable. A considerable improvement of the accuracy of potential has been observed for the case in which both energy and forces are used with respect to the one where only energy is used.
As reported in Ref. [180], the environment dependence of the NNP can effectively be reduced by the environment dependence of the NNP can effectively be reduced by

\[ R_c \]

extended beyond the cut-off radius \( R_c \) by including force fitting, which is necessary to keep energy and forces consistent. According to the comparisons of the performance of NNP1 and NNP3, the inclusion of the force fitting turned out to be crucial for developing NNPs applicable to large-scale MD simulations. Consequently, the code \( n2p2 \) instead of \( \text{aenet} \) has been used for constructing the potential applied in non-cumulative self-sputtering simulations on the beryllium surface in Ref. [4]. The MD simulations were achieved by an interface to the LAMMPS code.
3.3 Neural network approach

**The size of the training set**

Different sizes of training set were randomly chosen from the training data with 50,000 configurations mentioned above to fit various NNPs to study the NNP accuracy dependence on the training set size. The NNPs were employed in short MD simulations to generate new structures and the energies and forces of these structures were recalculated by DFT. Then the RMSE of energies and forces predicted by NNPs were calculated with respect to DFT data.

As seen in Figure 3.7, if we only choose 10 or 100 structures, both the energies and forces have a large error. For this specific training data, 500 configurations seem to be good enough and adding more structures from the data base does not improve the accuracy of the NNP.

![Figure 3.7: The root mean square error of energies and forces from NNPs in comparison to reference DFT data. These NNPs are fitted using different sizes of the training set as randomly chosen from the full data set.](image)

**Some beneficial settings of n2p2**

**A. Symmetry function groups**

In the n2p2 library [175] the symmetry function set and atomic NNs classes are specified for each element. To compute the total potential energy for one configuration, in the innermost loop the symmetry function set is calculated and implemented as input
for the corresponding NN. Subsequently, the atomic energy is fitted by element specific NNs, and then the total potential energy is obtained by summing up the atomic energy over all atoms in the configuration [175].

Calculating the symmetry function set is a computational demanding procedure, and the cost increases rapidly with increasing the size of training set and the number of element types. To split the symmetry function set into several symmetry function groups is a simple but effective way to reduce the computational cost [175]. The basic idea is to avoid repeated calculations for symmetry functions over the same neighbor atoms for different parameters ($\eta$, $\lambda$ and $\zeta$).

From the definition of angular symmetry functions in Eqn. (3.32), all the cut-off function terms $f_c(R_{ij})$, $f_c(R_{ik})$ and $f_c(R_{jk})$ share the same cut-off radius $R_c$, namely the same list of neighboring atoms. Therefore, it is not necessary to calculate the cut-off functions for each symmetry function. Meanwhile, the gradient of these cut-off functions required for calculating atomic forces (see Eqn. (3.34)) can also be reused. It is certainly efficient to group some symmetry functions together according to the element types of the neighboring atoms and the values of cut-off radius and other parameters. Part of angular symmetry function groups for Be as the central atom for the training set used to construct the NNP applied in the sputtering simulations on Be surface by H impacts are listed in Table 3.2.

As shown in Table 3.2, the cut-off functions are calculated for each group but not for each symmetry function. With the same value of $\eta$, the exponential term $e^{-\eta(R_{ij}^2+R_{ik}^2+R_{jk}^2)}$ can also be reused for different symmetry functions in the same group.

There are certainly many choices for symmetry function parameters. Refs. [189, 190] reported that the performance of describing the local chemical environment is not sensitive to the selected parameters but to the number and the type of symmetry functions. For a single chemical environment, about 50 symmetry functions with different types are normally used [179]. In Ref. [4], 53 symmetry functions including 9 radial, 24 angular and 20 angular wide functions were employed to develop a NNP for Be self-sputtering simulations.

### B. Multi-stream Kalman filter

The Kalman filter algorithm uses noisy data points observed over a whole time history to predict unknown variables. It is conceptually related to similar algorithms from Bayesian logics. The estimated variables tend to be more accurate than the ones that are based on a single observation alone. The algorithm works recursively and requires only an error
### 3.3 Neural network approach

<table>
<thead>
<tr>
<th>Element 1</th>
<th>Element 2</th>
<th>( \eta ) (bohr(^{-2}))</th>
<th>( \lambda )</th>
<th>( \zeta )</th>
<th>( R_c ) (bohr)</th>
<th>Exponential factor</th>
<th>Cut-off function</th>
</tr>
</thead>
<tbody>
<tr>
<td>Be</td>
<td>Be</td>
<td>0.2283</td>
<td>13.0</td>
<td></td>
<td></td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1 1.0</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 1.0</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1 4.0</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 4.0</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1 9.0</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 9.0</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>Be</td>
<td>H</td>
<td>0.2283</td>
<td>13.0</td>
<td></td>
<td></td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1 1.0</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 1.0</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1 4.0</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 4.0</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1 9.0</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 9.0</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td>H</td>
<td>H</td>
<td>0.2283</td>
<td>13.0</td>
<td></td>
<td></td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1 1.0</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 1.0</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1 4.0</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 4.0</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>-1 9.0</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>1 9.0</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

Table 3.2: Part of beryllium angular symmetry function groups for the H-Be system. Element 1 and element 2 denote the elements of the neighbor atoms of the central Be atom. The column “exponential factor” and “cut-off function” represent whether the exponential factor and cut-off function (as well as its gradient) in Eqn. (3.32) are needed to be recalculated, respectively.

Kalman filters have wide applications, especially navigation, aircraft and vehicle control. Since the original Kalman filter is limited to linear models, an extended Kalman filter [191] was soon developed for non-linear models. The extended Kalman filter was found to be applicable in the NN training process. In the application of NN training, the estimate state is described as the weight parameters vector \( \vec{\omega} \). The measurements containing covariance matrix and the last estimate state to estimate the current state.
noises are energies and forces [186]. The current state vector is updated recursively:

\[
\vec{\omega}(t+1) = \vec{\omega}(t) + \vec{K}(t)\vec{\xi}(t) \tag{3.35}
\]

where \(\vec{K}(t)\) is Kalman gain matrix and a detailed discussion about calculating this matrix can be found in Ref. [186]. \(\vec{\xi}(t)\) is the error vector, which is defined as:

\[
\vec{\xi}(t) = \vec{y}^{ref} - \vec{y}(t) \tag{3.36}
\]

where \(\vec{y}^{ref}\) is the vector consisting reference energies and forces, and \(\vec{y}(t)\) is the NN predictions of energy and forces with the current state vector \(\vec{\omega}(t)\), i.e. weights. The weights are optimized by minimizing the cost function \(\Gamma = \sum \vec{\xi}^T(t)\vec{\xi}(t)\) [186].

For a given configuration with index \(a\) consisting of \(N_a\) atoms, \(\vec{\xi}(t)\) would have \(3N_a+1\) components, \(3N_a\) of which are based on the atomic forces on three directions and only one of which is based on energy.

\[
\vec{\xi}(t) = \begin{bmatrix}
E^{ref}_a - E_a(\vec{\omega}(t)) \\
F^{ref}_{a1,x} - F_{a1,x}(\vec{\omega}(t)) \\
F^{ref}_{a1,y} - F_{a1,y}(\vec{\omega}(t)) \\
F^{ref}_{a1,z} - F_{a1,z}(\vec{\omega}(t)) \\
\vdots \\
F^{ref}_{aNa,z} - F_{aN_a,z}(\vec{\omega}(t))
\end{bmatrix} \tag{3.37}
\]

From Eqn.(3.37), one can expect inaccurate energy predicting results, if the error vector is populated by all the information from forces with equal weighting. One simple way to deal with this problem is to treat the contributions from energy and forces separately, and in such way the cost function described in Eqn. (3.38) is minimized [186].

\[
\Gamma = \sum_{a=1}^{N} (E^{ref}_a - E_a)^2 + \tau \sum_{a=1}^{N} \sum_{i=1}^{N_a} (F^{ref}_{ai} - \bar{F}_{ai})^2 \tag{3.38}
\]

Here, \(N\) is the number of the configurations in the training set and \(\tau\) is a free parameter that is introduced to control the influence of the forces.

C. Storing symmetry function values
In the \(n2p2\) library, one can choose whether to memorize the symmetry functions. Since the symmetry functions are the input of a NN, if they can be stored and reused, the
training speed can be significantly increased. The derivatives of the symmetry functions with respect to atomic positions can also be reused for calculating the atomic forces given in Eqn. (3.34). However, memorizing symmetry function results usually requires a large amount of main memory. A rough estimate of memory usage for different size of training sets for fitting the NNP for two elements (H and Be) is given in Table 3.3. Here, the NN input for H consists of 18 radial and 36 angular Behler-type symmetry functions, and the one for Be consists of 18 radial and 54 angular symmetry functions. The number of atoms for each configuration is about 100 atoms. The cut-off radius is chosen to be 13 Bohr, corresponding to an average of 127 neighbors per atom.

One can see that the memory usage increases lineally with increasing the number of configurations in the training set. Thanks to several nodes equipped with large main memory (512 GB) on LEO3E and LEO4 clusters in the university of Innsbruck, the training process with training sets including up to 20,000 configurations can be performed on these big memory nodes while memorizing symmetry functions.

<table>
<thead>
<tr>
<th>Number of configurations in the training set</th>
<th>Estimated memory usage (GB)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2422</td>
<td>60</td>
</tr>
<tr>
<td>4148</td>
<td>126</td>
</tr>
<tr>
<td>6321</td>
<td>155</td>
</tr>
<tr>
<td>8247</td>
<td>201</td>
</tr>
</tbody>
</table>

Table 3.3: An estimate of memory usage for training in the condition that symmetry functions are stored.
4 Summary of results

The scientific results achieved in this thesis are reported in the four peer reviewed publications [1]-[4], if not otherwise indicated below. This section gives an overview of them with examples of the results and an overview of their connections is given in Figure 4.1.

Figure 4.1: The flowchart of the connections between publications [1]-[4].
4.1 Performance of DFT functionals for neutral Be/W/H molecules

Due to the exposure of Be and W walls to the hot plasma containing hydrogen and its isotopes, the erosion of the plasma facing wall leads to the formation of Be and W hydrides and Be-W compounds. The electronic properties of these compounds and their reactions should be quantitatively investigated since they influence the capacities of the plasma facing surface and many reactor properties. For example, obtaining accurate geometrical parameters is an essential work for calculating the electron-impact ionization cross sections (EICs) of fusion relevant molecules [192].

The ground state electronic properties (bond lengths, atomization energies and vibrational frequencies) of a set of neutral Be/W/H molecules \( \text{Be}_n\text{H}_m, \text{W}_n, \text{W}_n\text{Be}_m, \text{and W}_n\text{H}_m \) with \( m + n \leq 4 \) have been studied by 16 DFT functionals from rungs 1 to 4 of Jacob’s ladder and their performance was compared with respect to \textit{ab initio} coupled cluster methods in Ref. [1].

The studied molecules cover both main group chemistry and transition metals. Their investigation is not only of interest for fusion research but also provides insight into the performance of density functionals. \( \omega \text{B97XD}, \) closely followed by B97D, M06, B3LYP and M11, show the best performance for the atomization energies, while three range-separated functionals (M11, \( \omega \text{B97XD} \) and HSEH1PBE) best reproduce bond lengths. In the case of predicting vibrational frequencies, the ranking sequence is HSEH1PBE, B3LYP and M11. The range-separated hybrid meta-GGA functional M11 was found to be accurate for all three properties for all studied molecules, with a mean absolute error (MAE) of 0.07 eV per atom for energy, an MAE of 0.007 \( \text{Å} \) for bond lengths and a mean absolute relative error (MARE) of 2.5% for vibrational frequencies.

The basis set and size dependence of the CCSD(T) energies and the influence of the core electrons on the energies of \( \text{W}_2 \) and W hydrides (\( \text{WH}, \text{WH}_2 \) and \( \text{WH}_3 \) were also discussed. We found that CCSD(T)/cc-pVQZ energies generally agree well with cc-pVQZ/cc-pV5Z extrapolations to the CBS limit. All-electron calculations show that core correlation energy contributes with 3-5% for \( \text{W}_2 \) and W hydrides in energies.

Furthermore, we also compared the ground state properties with PW DFT results which were not included in Ref. [1] but are shown in this section. We compared the results of PW calculations against Gaussian Type Orbital (GTO) based (cc-pVQZ) ones for the
4.1 Performance of DFT functionals for neutral Be/W/H molecules

PBE functional.

Plane wave basis sets are normally used for periodic systems. Nevertheless, a molecule in vacuum can be described by using a large supercell and a large number of plane waves and some properties of plane waves make such calculations popular. For example, PW calculations do not suffer from basis-set superposition errors. The results of both types of basis sets should of course be identical in the limit of a complete basis set, but the technical details of actual calculations cause deviations. Since both methods are routinely applied in material science calculations but hardly compared, we report these findings here.

The calculations employing a plane wave basis were carried out with the VASP program. For them, the size of the supercell was $12 \times 12 \times 12 \text{ Å}$, which has been carefully selected for convergence to avoid interactions between neighbor images (Figure 4.2). The Projector Augmented Wave (PAW) [193] method and the PBE [127] exchange-correlation functional describe the core and valence electrons. The same values of kinetic cut-off energy (550 eV), k-point mesh ($5 \times 5 \times 5$) and convergence criteria on the forces ($10^{-4}$ eV/Å) that have been tested and used in Refs. [2, 72] were used.

![Figure 4.2: Energy of the WH$_3$ molecule as a function of the size of the super cell. Here the horizontal coordinates $x_i$ means that the super cell size is $x_i \times x_i \times x_i$ Å$^3$.](image)

54
In Ref. [1], the CBS limit from CCSD(T) cc-pVQZ/cc-pV5Z extrapolations is used as the reference energy for comparing the performance of DFT functionals. The absolute value of deviation per atom of plane wave and PBEPBE functional with GTO with respect to the CBS limit ($E_{\text{ref}}$), defined as $|E - E_{\text{ref}}|$, is shown in Figure 4.3.

For most beryllium and tungsten hydrides, GTO and plane waves give similar performance with an error smaller or a little bit larger than 0.1 eV/atom. For pure Be molecules, both GTO and plane waves have very close but large errors. This is not unexpected. As is mentioned in Ref. [1], Be$_2$ is unsuitable to be treated with almost all the studied DFT functionals. The beryllium dimer, in spite of having only eight electrons, has to be treated with very accurate ab initio methods, like full CI [194] or couple cluster
methods with very big basis sets [195] for approaching experimental findings. For the tungsten dimer, WBe and Be$_3$H, GTO show a slightly better performance than plane waves.

Overall, the plane waves DFT calculations are comparable to the same PBE exchange and correlation functional with Gaussian orbitals. Due to its demanding computational efforts, plane waves DFT is not recommended for studying small molecules, but it shows great advantages in investigating solid or liquid systems. By applying periodic boundary conditions, the problems of an infinite system are converted into studying the limited atoms in a supercell with an appropriate size.

4.2 DFT study for surface defect energetics of Be, W and their alloys

We employed plane wave DFT calculations to calculate surface defect and surface binding energies for Be, W, Be$_2$W and Be$_{12}$W surfaces with various slabs in (110) and (111) directions [2]. The values of surface binding energy $E_{SBE}$ are in a wide range (4.45 ± 1.18 eV for beryllium and 8.54 ± 1.73 eV for tungsten), depending regularly on the number and types of neighbor atoms. The energetics for addition of a single atom on the surface slabs were also calculated, and these energetic data provides essential insight into the evaluation of deposition and adsorption processes. Concerning comparable structures, the values of adatom binding energies are smaller than or roughly equal to the $E_{SBE}$ values. This confirms that adatoms are more easily sputtered than surface atoms. In analogy to $E_{SBE}$, W adatoms on the Be$_{12}$W surface slabs are more weakly bound than on the Be$_2$W structures, and Be adatoms mostly follow the same trend. In terms of higher-order defects (surface steps), the binding energies of ledge atoms on surface steps have also been studied, and ledge atoms were found to be more stable than adatoms for comparable structures. The large differences of the studied energetic parameters for various surface slabs and directions mean that our DFT calculations are essential for understanding the surface stabilities, adatom and deposition process for PFM in the ITER environment.
4.3 H adsorption and penetration into the Be$_2$W surface

We have employed Kohn-Sham density functional theory to investigate hydrogen adsorption and penetration in some of PFM on an atomistic modelling level. An example is the hydrogen adsorption on a Be$_2$W(001) surface.

4.3.1 Details of the calculations

Kohn-Sham DFT was used to optimize geometries and to predict the ground state energies by using the VASP code [143, 144]. The PBE [127] functional was used to describe electron exchange-correlation interaction. A PW basis set with a cut-off energy $E_{\text{cut}}$ of 550 eV were employed. The $E_{\text{cut}}$ value has been tested and used in our previous work in Refs. [2, 72]. PAW potentials for W (with six valence electrons), Be (with two valence electrons) and H were taken from the VASP library. The Gamma-centered k-point mesh of 5×5×1 was employed. The first lowest layer of atoms was fixed while all other atoms were allowed to relax with a convergence criterion of $10^{-4}$ eV/Å on the forces during the geometry optimization processes.

4.3.2 Surface model

We created the stoichiometric systems as slabs for our surface investigations. As is discussed in Ref [2, 72], the surfaces of Be$_2$W alloy with (001), (110) and (111) orientations have various slabs. We chose slab D and slab A of the Be$_2$W(001) surface introduced in Ref. [72] as two example slabs to study the adsorption and retention process of H atoms. Side view of these two slabs are given in Figure 4.4. The thickness of the vacuum was chosen to be 13 Å, which is large enough to avoid the opposing surfaces disturbing each other. The super cell dimensions along the supercell vectors $\vec{a}, \vec{b}, \vec{z}$ are 8.9×8.9×26.7 Å$^3$, including 64 Be atoms and 32 W atoms. The crystal structure of Be$_2$W is hexagonal (space group P63/mmc). The $\vec{z}$ vector is perpendicular to the $\vec{a}, \vec{b}$ plane, and the angle between $\vec{a}$ and $\vec{b}$ vector is 120°. The lattice parameters for Be$_2$W optimized in Ref. [72] were used.
4.3 H adsorption and penetration into the Be$_2$W surface

4.3.3 Single hydrogen atom adsorption on slab D

Adsorption energies

For examination of hydrogen adsorption, five types of high-symmetry adsorption sites on slab D were tested as initial configurations for static relaxation of individual hydrogen atoms: two threefold hollow sites of three Be atoms in the surface layer (T1 and T2), one atop site (AT), one bridge site between two Be atoms (B1) and one bridge site between two W atoms (B2). These five adsorption sites along with the three topmost layers are shown in Figure 4.5.

The hydrogen atom was started from 2 Å above the five adsorption sites shown in Figure 4.5 and then the system was relaxed for geometry optimization. The adsorption energies of H are calculated as follows:

$$E_{ad} = E_{S+H} - E_S - \frac{1}{2}E_{H_2}$$  \hspace{1cm} (4.1)

where $E_{S+H}$ is the energy of the optimized surface with adsorbed hydrogen atoms, $E_S$ is the energy of the clean surface, and $E_{H_2} = -6.472$ eV is the DFT calculated energy of molecular H$_2$. The calculated adsorption energies of H atom on various adsorption sites are summarized in Table 4.1. $E_{ad,ZPE}$ means that Zero-point energy (ZPE) corrections have been included in the calculations of the adsorption energies.
4.3 H adsorption and penetration into the Be$_2$W surface

Figure 4.5: Five adsorption sites of the slab D of Be$_2$W(001) surface. The surface layer consisting of Be atoms (dark green), the first subsurface layer consisting of W atoms (gray), the second subsurface layer consisting of Be atoms (green), and H adsorption positions (orange) are shown.

<table>
<thead>
<tr>
<th>Adsorption sites</th>
<th>$E_{ad}$ (eV)</th>
<th>$E_{ad,ZPE}$ (eV)</th>
<th>$H \perp$ surf. (Å)</th>
<th>H-Be (Å)</th>
<th>H-W (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>-0.39</td>
<td>-0.28</td>
<td>1.46</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>-0.04</td>
<td>0.08</td>
<td>-0.62</td>
<td>1.63</td>
<td></td>
</tr>
<tr>
<td>AT</td>
<td>-0.51</td>
<td>-0.40</td>
<td>1.39</td>
<td>1.39</td>
<td></td>
</tr>
<tr>
<td>B1</td>
<td>-0.51</td>
<td>-0.38</td>
<td>1.07</td>
<td></td>
<td>1.81</td>
</tr>
<tr>
<td>B2</td>
<td>-0.58</td>
<td>-0.47</td>
<td>0.57</td>
<td>1.47</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.1: Adsorption energies of five adsorption sites on slab D of Be$_2$W surface. $H \perp$ surf. denotes the distance of H atom from the surface plane. There the negative value means that H atom is inside the surface. H-Be/W denotes the distance of H atom from its closest Be or W atoms.

As is given in Table 4.1, the bridge site B2 is the most energetically favorable adsorption site with the lowest $E_{ad,ZPE}$ of -0.47 eV and competes with AT and B1. Here, surface Be atoms are slightly shifted by 0.02 Å along the [00$\bar{1}$] direction. The H atom is 0.57 Å above the surface level, at equal distance (2.36 Å) from two neighboring W atoms, bonded to a Be neighbor with the bond length of 1.47 Å.

The AT site is slightly less stable than the B2 site, with a Be-H bond length of 1.39 Å. This agrees well with the length of the Be-H bond (1.37 Å) obtained by A. Allouche et al. [84], where an H atom was adsorbed in the atop site of a Be(0001) surface. However, in another work, the AT site on the Be(0001) surface was found to be unstable, and the H atom tended to approach threefold sites [86].

59
The H atom in the bridge site B1 with an adsorption energy $E_{ad,ZPE}$ around 0.1 eV higher than B2, is bonded to its nearest W neighbor with a bond length of 1.81 Å. Here, the H atom is at equal distance (2.37 Å) from two neighboring Be atoms.

It has also been pointed out in Ref. [84] that an H atom adsorbed on the Be(0001) surface is more likely to be bound to three equivalent Be atoms than right upon Be. In our study, the trend is opposite: the threefold site T1 with an $E_{ad,ZPE}$ of -0.28 eV is less stable than the atop site AT with an $E_{ad,ZPE}$ of -0.4 eV; The T2 site having a positive value of $E_{ad,ZPE}$ is not even stable. To a first explanation, Be atoms are much more close-packed in the pure Be(0001) surface with the distance of 2.27 Å than in the slab D of the Be$_2$W(001) surface with the distance of 4.46 Å. Additionally, the DFT calculated surface binding energy of 5.13 eV for the Be(0001) surface is larger than 4.08 eV for slab D of the Be$_2$W(001) surface [72].

For the T1 site, during the relaxation process the distances between H atom and its three neighboring Be atoms in the surface layer are too far to form a bond. H atom continues to approach the W atom in the first subsurface layer, and the minimum energy was found for a W-H distance of 1.79 Å.

The H atom at the T2 site, without the barrier of W atoms, can continue to drop down into the surface until it is bound to three Be atoms in the second subsurface layer. Here, the H atom is 0.62 Å below the surface slab, with a Be-H bond length of 1.63 Å. For H adsorbed in threefold site on a pure Be surface, the previous theoretical estimate for the Be-H bond length is 1.56 Å [86].

**Charge density distribution**

H adsorption on the substance surface leads to charge redistribution. The local charge density difference $\rho_{diff}$ is defined as:

$$\Delta \rho = \rho_{Be_2W+H} - \rho_{Be_2W} - \rho_H$$

In Eqn. (4.2) $\rho_{Be_2W+H}$ is the Charge Density (CD) for the optimized Be$_2$W surface with an adsorbed hydrogen atom, $\rho_{Be_2W}$ and $\rho_H$ are the CD of the substance and the adsorbed hydrogen atom separately. 3D images of CD differences and its integration over the $\vec{a} \vec{b}$ planes $\Delta \rho(z)$ of H adsorbed on the aforementioned five sites are given in Figure 4.6.

As is shown in Figure 4.6, the bonding between H atom and the surface influences
4.3 H adsorption and penetration into the Be$_2$W surface

Charge density difference for Be$_2$W-H system of different adsorption sites. For each image, the left is the charge difference $\Delta \rho(z)$ integrated over an $\vec{a}$, $\vec{b}$ plane. The origin of $Z$ is the surface layer. The position of the H atom is indicated by the red dash line. A positive value of $\Delta \rho(z)$ indicates an average electron accumulation. The Plot on the right shows 3 dimensional iso-surfaces of the electron-density difference at $\Delta \rho = \pm 0.002 \text{ e/}\AA^3$. Blue regions indicate electron density accumulation, and red regions indicate electrons depletion.

The electron density distribution up to 4 Å below the surface. The electrons accumulation regions (blue lobes) are centered on H atom and depletion regions are centered on its closest neighboring Be or W atoms.

For the adsorption sites of T2, AT, B1 and B1 with formation of Be-H bonds, the higher electron affinity of a H atom (0.75 eV) when compared with that of a Be atom (0 eV) leads to electron transfer from Be atoms to H. Although the electron affinity of H is very close to that of W, for the T1 site due to the presence of Be neighbors, a charge transfer to H occurs.

For the T2 site, the H atom is located inside the surface, bonded to three Be atoms in
the second subsurface layers. Strong repulsive interaction between electrons makes T2 cite not stable only due to ZPE effects. The electron depletion region is not only centered on the three bonded Be atoms and but also slightly above the surface plane. For the most stable adsorption cite B2, the charge redistribution takes place around the adsorbed H atom and its three neighborhoods (one Be and two tungsten atoms), which means that the H atom strongly interacted with these three atoms.

### 4.3.4 Single hydrogen atom adsorption on slab A

Three high-symmetry adsorption sites were tested as initial configurations for molecular static relaxation of individual hydrogen atoms on slab A of Be$_2$W(001): a threefold hollow site between W atoms in the surface layer (T$^*$), the atop site (AT$^*$), the bridge site between two W atoms (B$^*$). We insert a superscript “*” to distinguish slab D from slab A. The adsorption sites on slab A along with the two topmost layers are shown in Figure 4.7. The calculated adsorption energies of the H atom on these three adsorption sites are summarized in Table 4.2.

Likewise, the bridge site B$^*$ is the most stable site for a single H atom adsorbed on slab A of Be$_2$W (001) surface with the $E_{ad,ZPE}$ of -0.32 eV. The H atom is 0.40 Å above the surface level, at equal distance (2.36 Å) from two neighboring W atoms, bonded to a Be atom in the first subsurface layer with a bond length of 1.55 Å. The atop site AT$^*$ with the $E_{ad,ZPE}$ 0.2 eV higher than the B$^*$ site is slightly less stable. The W-H bond length was found to be 1.82 Å. The threefold site T$^*$ is energetically not stable, similar to the H adsorption on slab D. In contrast to the adsorption sites on slab D as mentioned above, adsorption sites on slab A are generally less stable for comparable structures.

<table>
<thead>
<tr>
<th>Adsorption sites</th>
<th>$E_{ad}$(eV)</th>
<th>$E_{ad,ZPE}$(eV)</th>
<th>H $\perp$ surf. (Å)</th>
<th>H-Be (Å)</th>
<th>H-W (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>AT$^*$</td>
<td>-0.23</td>
<td>-0.12</td>
<td>1.82</td>
<td></td>
<td></td>
</tr>
<tr>
<td>T$^*$</td>
<td>0.00</td>
<td>0.06</td>
<td>0.02</td>
<td>1.75</td>
<td>1.82</td>
</tr>
<tr>
<td>B$^*$</td>
<td>-0.42</td>
<td>-0.32</td>
<td>0.40</td>
<td>1.55</td>
<td></td>
</tr>
</tbody>
</table>

Table 4.2: Adsorption energies of three adsorption sites on slab A of Be$_2$W surface. H $\perp$ surf. denotes the distance of H atom from the surface plane. The negative value means H atom is inside the surface. H-Be(W) denotes the distance of H atom from its closest Be or W atoms.
4.3 H adsorption and penetration into the Be$_2$W surface

![Figure 4.7: Three adsorption sites on slab A of Be$_2$W(001) surface. In the picture, the surface layer consisting of W atoms (gray), the subsurface layer consisting of Be atoms (green), and the H adsorption positions (orange) are shown.]

4.3.5 Penetration of H atoms into the Be$_2$W subsurface

The adsorbed H atom must overcome an energy barrier that prevents H from penetrating into the surface. We found a local stable site in the subsurface region denoted as subsurface site as shown in Figure 4.8. The H atom was adsorbed in the center of two W atoms in the third subsurface layer, bonded to two Be neighbors with the bond length of 1.49 Å and 1.52 Å respectively and to one W neighbor with the distance of 1.80 Å. It is worth noting that the adsorbed H atom and its bonded neighbors are in the same plane.

The penetration barrier via the most stable adsorption site B2 into the subsurface site was calculated using the climbing image nudged elastic band (CI-NEB) method [196]. The calculated penetration energies and migration path are shown in Figure 4.9. The retention barrier is calculated to be 1.35 eV, which is 0.7 eV lower than the barrier of H atom diffusing from the reconstructed W(100) surface to the bulk obtained by DFT calculations (see Ref. [94]). Thus the probability of the retention in the Be$_2$W surface maybe much higher than those in the pure W surface. On the other hand, the H atom is easily desorbed from the subsurface site to the surface, with the desorption barrier of 0.1 eV. The saddle point is found to be the threefold hollow site of three Be atoms in the second subsurface layer, equally bonding to the Be atoms with the bond length of 1.38 Å. This adsorption site is similar to the T2 site shown in Figure 4.5, but with the H atom
being deeper penetrated into the surface. The $E_{ad}$ of the saddle point is 0.8 eV higher than that of T2. The Be atoms in the second subsurface layer are the obstacles to the retention of H atom.

The diffusion rate of a solute atom in solid materials generally follows an Arrhenius equation

$$D = D_0 e^{-\frac{E_D}{kT}}$$  \hspace{1cm} (4.3)

where $D_0$ is the pre-exponential factor, $T$ is temperature and $k$ is the Boltzmann constant and $E_D$ is the activation energy of diffusion. For the migration path shown in Figure 4.9(a), the value of $E_D$ is 1.35 eV. Based on the diffusion theory reported by Wert and Zener [197, 198], the diffusion rate can be written as

$$D \approx \frac{1}{2} \lambda^2 \Gamma$$  \hspace{1cm} (4.4)

with $\lambda$ being the jump length taken equal to 5.64 Å from Figure 4.9(b) and $\Gamma$ being the jump rate. In a purely diffusion process without any chemical reaction at the saddle point, and if all vibrations at both the ground and activation state are harmonic, Vineyard
4.3 H adsorption and penetration into the Be$_2$W surface

Figure 4.9: (a) Migration path of H adsorption from the surface bridge site B2 to the subsurface site as obtained with CI-NEB method. (b) Potential energy pathway as a function of H position on the migration path shown in (a).

[199] reported the jump rate to be

$$\Gamma = \frac{\prod_{i}^{3N} v_i}{\prod_{j}^{3N-1} v_j} e^{-\frac{E_D}{kT}}$$  \hspace{1cm} (4.5)$$

where $v_i$ and $v_j$ are the real normal vibration modes on the ground state and saddle point respectively. It is worth remembering that the potential energy surface yields an imaginary vibrational frequency at the saddle point. This leads to one less real vibrational mode at the activation state than at the ground state. Combing Eqns. (4.4) and (4.5), the diffusion rate for single H atom diffusion from the B2 site into Be$_2$W surface is calculated to be $D=1.6 \times 10^{-6} e^{-1.35 \text{ eV}/kT} \text{ m}^2/\text{s}$.

We investigated hydrogen adsorption on various adsorption sites on slab D and slab A of a Be$_2$W(001) surface using density functional theory. We found the bridge site between two W atoms B2 on slab D is the most energetically favorable site. The charge density difference analysis indicated a strong interaction between the adsorbed H atom and its neighboring one Be and two W atoms. The adsorption sites on slab A are overall less stable than the comparable sites on slab D. Electron depletion is centered above on the neighboring atoms of H rather than on H atom itself from a previous study on a pure W surface [200].

The H penetration process via the B2 site to a subsurface site was also studied. The activation barrier calculated using the CI-NEB method was 1.35 eV, and the diffusion rate was $1.6 \times 10^{-6} e^{-1.35 \text{ eV}/kT} \text{ m}^2/\text{s}$. A comparison with previous estimates by Heinona and Ahlgren gives the barrier of H atom diffusion via a reconstructed W(100)
4.4 Neural Network potential energy functions and their application to the self-sputtering of a Be surface

Ref. [3] describes an interface between the NN library aenet 2.0.3 and the widely-used MD code DL_POLY 4.08 and its application to a liquid water system. It showed that the constructed NNP was capable of accurately reproducing the dynamics of water molecules that adhere to the analytical force field SPC/E subjected to thermodynamic results. However, in the aenet 2.0.3 library, only the total energy of configurations but no atomic forces are used to fit the NNPs. Based on comparisons of the accuracy of NNPs given in Section 3.3.3, we found that force fitting is essential for developing NNPs applied in more complex MD simulations. Therefore, the n2p2 library in which both energies and forces are used in the fitting process was employed to develop NNPs that are applied for self-sputtering simulations of the Be surface in Ref. [4].

In Ref. [4], we have implemented an iterative refinement process to generate a well-balanced and complete training set of total energies and atomic forces. The training set was used to fit a NNP, which was successfully applied in reactive self-sputtering simulations for the beryllium surface at various incident energies. The obtained sputtering yield of 5.6% agrees perfectly with a yield of 5% obtained from full ab initio MD result at 100 eV, and the NNP simulations is more than two orders of magnitude faster than the ab initio MD simulation. At lower incident energies, NNP-based results compare well with earlier calculations based on pair potentials and bond-order potentials. The NNP developed in Ref. [4] can be used on larger time and length-scales and enables statistics similar to what is accessible by conventional force fields and reaching beyond the possibilities of ab initio MD simulations. Since there is no conceptual difference in treating more complicated systems, this approach may become the preferred one in developing potential energy functions for ternary systems, where the construction of a conventional potential becomes cumbersome.

The sputtering yield for the Be(0001) surface at 100 eV incident energy from ab initio MD
4.5 Development of NNP for Be sputtering simulation by D impacts

Based on the work in Ref. [4], we continued to develop potential energy functions using the neural network approach for sputtering simulations from a Be(0001) surface by energetic D impacts. The results are detailed below.

4.5.1 Training data generation and refinement

The static DFT and the Born-Oppenheimer \textit{ab initio} MD simulations to generate and to refine training data were performed using the VASP code. The PBE functional was used to describe electronic exchange-correlation interactions. A plane wave basis set with cut-off energy of 350 eV was used. The initial training set was generated by performing \textit{ab initio} MD simulations on a Be(0001) surface with 96 atoms. The geometry was first optimized, and then equilibrated within the NVT ensemble. The perpendicular impacts of D atoms at various incident energies were simulated. We used the same values of parameters (convergence criteria for geometry optimization, equilibration time, the energy of D impact, etc.) as used in Ref. [4] but a much smaller integration step of 0.1 fs.

We have implemented an iterative refinement protocol as shown in Figure 3.4 to refine the training data. With several refinement processes, the final reference data set consists of 3035 configurations containing 96 Be atoms and one D atom each. 2723 Total energies and 792 291 forces are used to train the final NNP and the remaining data is part of the test set which is used to validate the potential and to avoid overfitting. With 60 training steps, the RMSE in the test set converged to 1.7 meV/atom for energies and 0.18 eV/Å for atomic forces, close to the corresponding values in the training set (0.6 meV/atom for energies and 0.26 eV/Å for atomic forces). Overfitting was not observed in the fitting process.

The correlation of NNP and DFT energies per atom and atomic forces are shown in
4.5 Development of NNP for Be sputtering simulation by D impacts

Figure 4.10. Only the $x$-component is plotted in the right panel since $y$ and $z$ components show similar trends. We found that the values of NNP and DFT both for energies per atom and for atomic forces are almost perfectly correlated.

4.5.2 Non-cumulative sputtering simulations of a Be(0001) surface by D impacts using NNP

In Ref. [4] we show that although only small systems were used to train the NNP, it can be applied in reactive sputtering simulations for a larger system. The super cell was created by relaxing a Be(0001) surface, which consists of 490 atoms with a size of 15.6×15.6×30.1 Å. After the relaxation, the structure was equilibrated for 2 ps at 300 K within NVT ensemble. Then the surface was irradiated by a single D impact with kinetic energies of 10, 20, 50, 75 and 100 eV. The D atom was initially placed 5 Å above the surface and its x and y-coordinates were randomly shifted. Here, only perpendicular impacts were simulated. 2000 separate MD trajectories for 120 fs were performed with integration time step of 0.1 fs at high incident energies (50 eV, 75 eV and 100 eV), and 5000 separate MD trajectories were performed for low incident energies (10 eV and 20 eV). On condition that the decision of an observed sputtering event cannot be made within 120 fs, further 120 fs integration time were added to the corresponding trajectories. At these incident energies, the D atom has a very high initial velocity (∼1.0 Å/fs at 100
4.5 Development of NNP for Be sputtering simulation by D impacts

eV), and some projectiles were found to go through the super cell. Those projectiles were considered as retained in the bulk.

The sputtering yield, reflection rate (reflection of D atoms per total impacts) and retention rate (retention of D atoms per total impacts) are summarized in Table 4.3. They are compared with other MD simulations from Björkas et al. [201] and Sukuba et al. [202] where a bond-order potential was employed in Figure 4.11. The static error of the sputtering yield was estimated by $\Delta = 2\sigma/\sqrt{N}$, and $\sigma$ is the standard deviation obtained by assuming a Bernoulli distribution of $N$ trajectories.

<table>
<thead>
<tr>
<th>Impact energies (eV)</th>
<th>10</th>
<th>20</th>
<th>50</th>
<th>75</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beryllium sputtering yield</td>
<td>0.0004</td>
<td>0.0018</td>
<td>0.031</td>
<td>0.032</td>
<td>0.0325</td>
</tr>
<tr>
<td></td>
<td>(0.0006)</td>
<td>(0.001)</td>
<td>(0.008)</td>
<td>(0.008)</td>
<td>(0.008)</td>
</tr>
<tr>
<td>D reflection rate</td>
<td>15%</td>
<td>19%</td>
<td>18%</td>
<td>14%</td>
<td>12%</td>
</tr>
<tr>
<td>D retention rate</td>
<td>85%</td>
<td>81%</td>
<td>82%</td>
<td>86%</td>
<td>88%</td>
</tr>
</tbody>
</table>

Table 4.3: Sputtering yield (including static error), reflection rate and retention rate estimation for the Be(0001) surface irradiated by D impacts obtained from NNP at various incident energies.

Figure 4.11: Dependence of sputtering yield for a Be(0001) surface irradiated by non-cumulative D impact with various incident energy. The results of MD simulation by Björkas et al. [201] and by Sukuba et al. [202] are included for comparison.
The probability of reflection at 100 eV incident energy is 12%, agreeing with the simulation result (11%) from Björkas and Nordlund [201]. As the incident energy decreases, the reflection rate increases from 12% (100 eV) to 19% (20 eV). The sputtering yield is around 0.03 for incident energies above or equal to 50 eV, while it decreases rapidly to 0.002 for lower D energy. The sputtering threshold energy $E_{th}$ defined here by no sputtering yield occurring in 5000 trajectories lies closely below 10 eV, in a good agreement with previous estimate of 9.51 eV from literature [203].

We used the parameters $E_{th}=9.51$ eV, $q=0.10$, $\mu =1.99$ and $\lambda = 1.76$ from Ref. [203] to fit the simulated sputtering yields to the Eckstein formula, which is also included in Figure 4.11. At high incident energies (50, 75 and 100 eV), our result compare very well with the values from Björkas et al. [201]. At 20 eV, the NNP based sputtering yield is lower than the value from the bond-order potential [201]. The reason for this discrepancy is that an additional BeD chemical sputtering was involved at 20 eV in Ref. [201], which was not significant at higher energies, while in our simulations, we have not encountered chemical sputtering events. For even lower incident energy (10 eV), the yield obtained from NNP agrees with that from Sukuba et al. [202]
5 Outlook on future research

Many interesting research branches could not be followed yet, and here I just mention a few research directions.

Sun, Ruzsinszky and Perdew proposed a new semi-local density functional: strongly constrained and appropriately normed (SCAN) in 2015, which has achieved a major improvement over PBE for lattice constants of solids. SCAN is accurate in describing van der Waals interactions and in particular also rare-gas atoms [204]. It is therefore of interest to understand its performance in describing electronic properties of small Be/W/H molecules and Be/W surfaces.

The interaction between low energy ions and PFM is important for ITER operation, but comparable experiments are difficult to undertake and usually only give a final average picture over multiple campaigns. Thus, atomistic modeling plays an important role in understanding the mechanisms for predicting the PFM life time, T-retention, dust production and plasma contamination. In Ref. [4], we showed the applicability of predicting the Be self-sputtering yield with a machine learning approach. The technique will be further developed for the ternary H/Be/W system. A systematic study of the influence of the surface temperature, the incoming angle and the impact energy of H isotopes on the erosion properties of PFM as well as H retention in the surfaces shall be performed via MD simulations using neural network potentials. The sputtering yields and H isotopes retention rates of the Be/W alloys are not known yet. The effects of the formation of alloys on the erosion and the H retention properties, especially compared to pure PFM, shall be discussed. Quantum effects of H isotopes with very light mass at low energies cannot be completely ignored [205], while in MD simulations, so far we treated H classically. Path-integral molecular dynamics (PIMD) simulations [206] could shed light on these effects for low energy impacts.
6 Conferences

Besides the paper contributions, the author has also made oral or poster contributions to several international conferences and winter schools.


- DK CIM winter school in Obergurgl, Austria 2017; Poster

- “On the way towards accurate atomistic simulations of surfaces of Be, W and their alloys”, 4th Fusion@ÖAW Fusion Day, Leoben, Austria 2017; Oral talk.

- 27th IEEE Symposium On Fusion Engineering, Shanghai, China 2017; Poster.

- “Surface Properties of Be, W and Their alloys”, EUROfusion Joint SOL and PSI Modelling Activities, Aix-Marseille University, France 2018; Oral talk.

- DK CIM winter school in Obergurgl, Austria 2018; Poster

- “Developing the potential energy surface (PES) for BeWH using artificial neural network (ANN)”, 5th Fusion@ÖAW Fusion Day, Vienna, Austria, 2018; Oral talk.
Bibliography


Bibliography


surfaces: A prototype study for copper,” *Phys. Rev. B*, vol. 85, p. 045439, Jan
2012.

tials for atomistic materials simulations: Performance for TiO$_2$,” *Computational

interpolation of atomic energies in compositions with many species,” *Phys. Rev. B*,

[184] N. Artrith, A. Urban, Y. Wang, and G. Ceder, “Atomic-scale factors that control the
rate capability of nanostructured amorphous Si for high-energy-density batteries,”
2019.

of amorphous lixsi using machine-learning-assisted sampling with an evolutionary

[186] A. Singraber, T. Morawietz, J. Behler, and C. Dellago, “Parallel Multistream Train-

ters,” *Journal of the Society for Industrial and Applied Mathematics*, vol. 11, no. 2,

[188] K. LEVENBERG, “A method for the solution of certain non-linear problems in

and reference configurations for machine-learning potentials,” *The Journal of

[190] M. Gastegger, L. Schwiedrzik, M. Bittermann, F. Berzenyi, and P. Marquetand,
“wACSF—Weighted atom-centered symmetry functions as descriptors in machine
learning potentials,” *The Journal of Chemical Physics*, vol. 148, no. 24, p. 241709,
2018.


Performance of DFT functionals for properties of small molecules containing beryllium, tungsten and hydrogen

Lei Chen¹, Daniel Süß², Ivan Sukuba¹,², Michael Schauperl¹,³, Michael Probst³, Thana Maihom⁴, Alexander Kaiser¹,⁵

¹ University of Innsbruck, Institute of Ion Physics and Applied Physics, 6020 Innsbruck, Austria
² Department of Nuclear Physics and Biophysics, Comenius University, SK-84248 Bratislava, Slovakia
³ Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92093-0736, United States
⁴ Department of Chemistry, Faculty of Liberal Arts and Science, Kasetsart University, Kamphaeng Saen Campus, Nakhon Pathom 73140, Thailand

ARTICLE INFO

Keywords:
DFT functional comparison
Beryllium
Tungsten
Hydrogen
Coupled clusters
Complete basis set limit
Core correlation

ABSTRACT

A comparative study of bond lengths, atomization energies and vibrational frequencies for a set of neutral molecules Beₙ, BeHₙ, Wₙ, WBeₙ, and W₂Hₙ with m + n ≤ 4 in their ground state is presented. We compare 16 density functionals chosen from rungs 1 to 4 of Jacob's ladder and Hartree–Fock results to Coupled Cluster calculations. For atomization energies, ωB97XD, closely followed by B97D, M06, B3LYP and M11, shows the best performance while M11, ωB97XD and HSEH1PBE are most faithful in reproducing bond lengths. The ranking sequence for predicting vibrational frequencies is HSEH1PBE, B3LYP and M11. The range-separated hybrid mega-GGA functional M11 stands out to be accurate for all three properties for the studied Be/H/W molecules.

The basis set and size dependence of the CCSD(T) energies and the influence of inner core electrons on the atomization energies of several tungsten-containing molecules are also discussed. CCSD(T)/cc-pVQZ energies are already close to the complete basis set limit from a cc-pVQZ/cc-pV5Z extrapolation level. Core correlation contributes with 3–5% for W₂ and W hydrides in energies. The comparison of DFT functionals in this particular setting and the set of molecules allows to evaluate the accuracy of atomistic plasma-wall interaction studies for the ITER environment, where Be, W, and H/D/T interact with each other.

1. Introduction

The incentive of this study stems from plasma-surface interactions that can occur in thermonuclear devices. In particular, the International Thermonuclear Experimental Reactor (ITER) will employ beryllium (Be) and tungsten (W) as the armor of the first wall and divertor components, respectively [1]. Both of them will face the hot magnetically confined plasma composed primarily of hydrogen isotopes and helium [2]. Additionally, under the irradiation of 14 MeV neutrons from D-T fusion reactions, H can be produced in the plasma facing surface by nuclear transmutation (n,p) reactions [3]. Chemical erosion and sputtering of Be walls can happen, for example, due to the formation of beryllium deuterides [4], i.e. BeD, has been observed in the JET tokamak [5] as well as in a linear plasma device [6]. Atomic hydrogen trapped on the W(001) surface was confirmed by electron-energy-loss spectroscopy in 1977 [7], and the adsorbed H atoms can diffuse, accumulate, combine and finally lead to bubble formation that can significantly degrade the thermal–mechanical properties of W, which has been intensively studied and summarized in Ref. [8]. The structures, energetics and vibrational dynamics of small W hydrides that can provide fundamental insight into the reactivity of W and H, have been studied using ab-initio [9–12] calculations and neon matrix infrared spectra experiments [13]. In 2005, Juslin et al. [14] used a bond-order potential (BOP) for the ternary W–C–H system to reproduce energetics and structures of small W hydrides, in good agreement with results in Ref. [9–13]. On the other side, atomic beryllium eroded from the first wall may be transported, deposited and then adsorbed on tungsten surfaces [15], and a further penetration into tungsten bulk leads to the formation of intermetallic compounds that have been detected both in laboratory [16–20] and Tokamak [21] experiments. A recent test indicated that the Be–W mixed layer can cause a weakened resistance to edge localized modes (ELMs) like transient loads [22]. Furthermore, a DFT [23–26] calculation by Allouche, et al. showed a higher probability of trapping hydrogen in vacancies in the BeₓW bulk
than in pure beryllium [27]. In a previous study we derived geometries, atomization energies and ionization thresholds of BeₙW clusters (n ≤ 12) at the B3LYP/def2-TZVP level of theory in a study focused on their electron impact ionization cross sections [28].

The variety of compounds and the reactions between them should be quantitatively modelled since they influence plasma temperature, electron temperature, surface capacities and many reactor properties. This raises the question which quantum chemical method or density functional poses a consistently good compromise between feasibility and quality, if any. Therefore, we present a comparative study of equilibrium geometries, atomization energies and vibrational frequencies for the neutral molecules Beₙ, BeₙHₙ, Wₙ, WₙBeₙ, and WₙHₙ (m + n ≤ 4) with the aim to assess the accuracy of density functionals which are often employed in studies encompassing larger and also dynamic parts of the reactor environment. It is worth noting that the DFT functional comparison is valid for gas-phase calculations, but may not be transferable to periodic calculations of the metals.

2. Computational methods

We investigated the performance of 16 density functionals. Those were (a) the local spin-density approximation (LSDA) functional SVWN [29]; (b) the general gradient approximation (GGA) functionals BP86 [30,31], PW91PW91 [32–36], B97D [37,38], B97D3 [37,39], BLYP [30,40,41], PBE0 [42,43], and HCTH [44–46]; (c) the meta-GGA (M06L [47], TPSSPP [48]); (d) H-meta-GGA: hybrid GGA (B3LYP [29,40,49,50]) and hybrid meta-GGA (MO6 [51]); (e) RSH-meta-GGA: range-separated hybrid (RSH) GGA (oB97XD [37,38,52–54], HSEH1PBE [55–61], CAM-B3LYP [62]) and RSH meta-GGA (M11 [63]). Functionals that include empirical dispersion correction contain the keyword character ‘D’, i.e. B97D, B97D3 and ωB97XD.

Since experimental data is not available for most of the studied Be/H/W molecules, it is highly desirable to have accurate theoretical values as reference. For our set of molecules, CCSD(T) [64,65] calculations have to serve this purpose. We also compared the CCSD(T) data to the scarce available experimental data. The CCSD(T) calculations were performed with the cc-pVQZ basis set with an effective core potential (ECP) [66] for W. Extrapolations to the complete basis set (CBS) limit were performed for single point energies and the effect of the core-correlation was estimated by comparison of the ECP calculations with all electron calculations for the inner shells of W. The ORCA program was utilized for CBS extrapolations and all-electron calculations [67]. For the DFT calculations, various basis sets were compared on the functional PBE0PBE. DFT calculations were performed with the GAUSSIAN 16 software [68] for all DFT functionals.

3. CCSD(T) calculations

For the purpose outlined above it is crucial to have reference values as accurate as possible. The geometries of the studied molecules were optimized at the CCSD(T) level of theory with a widely used correlation-consistent polarized valence quadruple-zeta (cc-pVQZ) basis set, where ECPs [66] were included for tungsten. Calculated data on atomization energies including ZPE corrections are listed in Table 1 to explain the factor of two between our value and the experimental result.

Optimized geometries of the molecules at the CCSD(T) level are shown in Fig. S1 and Table S1 in the supplementary material. The ground-state bond lengths of BeH and BeH₂ calculated by CCSD(T) are 1.346 and 1.330 Å, respectively, which agrees very well with experimental values (1.342 Å for BeH [73] and 1.326 Å for BeH₂ [74]).

In terms of harmonic vibrational frequencies, the Be–H stretching mode of BeH is 2054.5 cm⁻¹. Somewhat unexpected, this is lower than the experimental value of 2061 cm⁻¹ [73]. The BeH₂ molecule is linear in its ground state (shown in Fig. S1), with vibrational frequencies of 712.2 (bending ν₁), 2045.4 (symmetric stretching ν₂) and 2254.0 (antisymmetric stretching ν₃) cm⁻¹, which deviates by a factor of about 0.98 and 0.96 respectively from the infrared spectroscopy results (697.9 cm⁻¹ for ν₁, 2159.1 cm⁻¹ for ν₂) by Tague and Andrews [75]. The vibrational frequency of diatomic WH is predicted to be 1933.5 cm⁻¹, requiring a 0.96 scale factor to fit the neon matrix measurement (1860 cm⁻¹) [13].

3.1. Complete basis set limit

The basis set dependence of the CCSD(T) energies of the 15 molecules from Table 1 was analyzed by employing the three types of basis sets def2, cc and aug-cc. The CCSD(T) energies were then extrapolated to the complete basis set (CBS) limit using the procedure described in Ref. [76]. The geometries of the molecules were taken at the CCSD(T)/cc-pVQZ optimized ones and were not reoptimized.

This extrapolation is described in Eqs. (1)–(3) [77,78].

$$E_{CC}^{X} = E_{CC}^{CBS} + A \exp(-\alpha \sqrt{X})$$  
(1)

$$E_{corr}^{X} = \frac{X P_{corr}^{X} - Y P_{corr}^{Y}}{X^{\beta} - Y^{\beta}}$$  
(2)

$$E_{CBS}^{X} = E_{CC}^{CBS} + E_{corr}^{CBS}$$  
(3)

Here, $E_{CC}^{X}$, $E_{corr}^{X}$ and $E_{CBS}^{X}$ are HF, correlation and total energies at the CBS limit, respectively. X and Y denote the cardinal number that is 3 for triple-zeta and 4 for quadruple-zeta, etc. The values of the constant α and β for different basis types and extrapolations have been determined in Ref. [76]. In Eq. (1), A and $E_{CC}^{CBS}$ can be obtained by using the SCF energy from triple- and quadruple-zeta basis sets as a two-point extrapolation. Similarly, the correlation energies $E_{corr}^{CBS}$ and $E_{corr}^{CBS}$ can be used to get $E_{corr}^{CBS}$ from Eq. (2).

Results from the CBS extrapolations including ZPE corrections are

<table>
<thead>
<tr>
<th>Molecule</th>
<th>Spin multiplicity</th>
<th>ZPE (eV)</th>
<th>CCSD(T) (eV)</th>
<th>Exp. (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H₂</td>
<td>1</td>
<td>0.27</td>
<td>4.46</td>
<td>4.48 (69)</td>
</tr>
<tr>
<td>Be₂H₂</td>
<td>1</td>
<td>0.01</td>
<td>0.05</td>
<td>0.10 (70)</td>
</tr>
<tr>
<td>Be₃H₃</td>
<td>1</td>
<td>0.08</td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td>Be₄H₄</td>
<td>1</td>
<td>0.21</td>
<td>3.47</td>
<td></td>
</tr>
<tr>
<td>Be₅H₅</td>
<td>2</td>
<td>1.03</td>
<td>2.03 (69)</td>
<td></td>
</tr>
<tr>
<td>Be₆H₆</td>
<td>2</td>
<td>0.31</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Be₇H₇</td>
<td>2</td>
<td>0.46</td>
<td>3.33</td>
<td></td>
</tr>
<tr>
<td>Be₈H₈</td>
<td>2</td>
<td>0.22</td>
<td>3.19</td>
<td></td>
</tr>
<tr>
<td>Be₉H₉</td>
<td>1</td>
<td>0.42</td>
<td>7.11</td>
<td></td>
</tr>
<tr>
<td>Be₁₀H₁₀</td>
<td>2</td>
<td>0.31</td>
<td>3.66</td>
<td></td>
</tr>
<tr>
<td>Be₁₁H₁₁</td>
<td>1</td>
<td>0.02</td>
<td>4.99</td>
<td>5 ± 1 (71)</td>
</tr>
<tr>
<td>WHe</td>
<td>7</td>
<td>0.03</td>
<td>1.10</td>
<td></td>
</tr>
<tr>
<td>WH₂</td>
<td>6</td>
<td>0.12</td>
<td>2.77</td>
<td></td>
</tr>
<tr>
<td>WH₃</td>
<td>5</td>
<td>0.28</td>
<td>5.56</td>
<td></td>
</tr>
<tr>
<td>WH₄</td>
<td>4</td>
<td>0.48</td>
<td>8.41</td>
<td></td>
</tr>
</tbody>
</table>
Table 2

<table>
<thead>
<tr>
<th>Type of basis</th>
<th>CBS extrapolations (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>def2-tzVPP</td>
</tr>
<tr>
<td>X/Y</td>
<td></td>
</tr>
<tr>
<td>H2</td>
<td>4.48</td>
</tr>
<tr>
<td>Be2</td>
<td>0.08</td>
</tr>
<tr>
<td>Be3</td>
<td>1.08</td>
</tr>
<tr>
<td>BeH</td>
<td>2.1</td>
</tr>
<tr>
<td>BeH2</td>
<td>6.17</td>
</tr>
<tr>
<td>BeH3</td>
<td>6.49</td>
</tr>
<tr>
<td>Be4</td>
<td>1.35</td>
</tr>
<tr>
<td>Be5</td>
<td>4.35</td>
</tr>
<tr>
<td>Be6</td>
<td>7.35</td>
</tr>
<tr>
<td>Be7</td>
<td>4.2</td>
</tr>
<tr>
<td>W2</td>
<td>4.85</td>
</tr>
<tr>
<td>WH</td>
<td>1.11</td>
</tr>
<tr>
<td>WH2</td>
<td>2.82</td>
</tr>
<tr>
<td>WH3</td>
<td>5.64</td>
</tr>
<tr>
<td>WH4</td>
<td>8.52</td>
</tr>
</tbody>
</table>

Atomization energies (eV) including ZPE from extrapolation to the complete basis set limit. X and Y are the corresponding cardinal numbers where 3/4 for example denotes an extrapolation from triple and quadruple zeta to the CBS.

Atomization energies at the CCSD(T)/cc-pVQZ level are shown for comparison in the rightmost column.

From Table 2, we note that the basis set size increases the extrapolated atomization energy for Be/H complexes (Hn, Be, and BeH) considerably by about 0.0 to 0.12 eV per atom when going from 3/4 to 4/5 with the cc-pVnZ basis, while it shows very small influence on W complexes (Wn, WnHm, and WnBe). Augmenting the basis set with small-exponent functions turns out to be less important and leads to slightly larger ones for W complexes (except for WBe) by about 0.00–0.04 eV per atom only. The def2-nZVP basis sets give similar results to cc-pVnZ for all molecules except for W and Be4 with a difference larger than 0.1 eV per atom. Overall, the CBS limit from extrapolations at the cc-pVQZ/cc-pV5Z level compares rather well with the CCSD(T)/cc-pVQZ and experimental energies and the extrapolated CBS limit will further be used as the reference energies in the comparison of DFT functionals in Section 6.

3.2. All-electron calculations

Our CCSD(T) calculations with aug-cc basis sets employ an ECP for tungsten. To estimate inaccuracies from treating the inner electrons this way, we calculated the energies of W2 and W hydrides using the all-electron basis set SARC as proposed in Ref. [79] for W and the def2tZVP basis set for H. This triple-zeta SARC basis set has been developed especially for transition metals Hf–Hg [79]. We did not perform a CBS-limit extrapolation due to the lack of fitted alpha and beta values (see Eqs. (1) and (2)).

Relativistic effects were taken into account by using the zeroth-order regular approximation (ZORA) Hamiltonian in the all-electron calculations. This Hamiltonian is scalar-relativistic and does not include spin-orbit coupling. Here we use again structures from CCSD(T)/cc-pVQZ geometry optimizations. The atomization energies are 5.39 eV, 2.78 eV, 5.65 eV and 8.59 eV for W2, WH, WH2 and WH3, respectively without including ZPE. When comparing the values from cc-pVnZ 4/5 extrapolations to the all electron calculations, the CBS extrapolations for W hydrides show larger atomization energies than the all-electron results with relative differences within 5%, only for W5 we obtain a smaller atomization energy by 3%. These differences can be primarily attributed to the core-correlation, since the difference in relativistic effects between the ZORA Hamiltonian and the ECP should be smaller [66,80].

4. Basis set dependence of the DFT calculations

Atomization energies have been calculated with the PBE/BE func-
tional and the three basis sets cc-pVTZ, cc-pVQZ, and cc-pV5Z with respect to cc-pVQZ basis sets.

The absolute value of relative deviation of the two smaller basis sets from cc-pVQZ (Eref), defined as \(\frac{E - E_{\text{ref}}}{E_{\text{ref}}}\), is shown in Fig. 1. The convergence from triple, quadruple to quintuple zeta values for all W containing molecules is rather regular. For pure Be and beryllium hydrides, the basis set size has only a rather small influence below 1% but interestingly, the quadruple-zeta basis gives similar or slightly bigger deviation for some of the molecules. The quadruple-zeta results are within 1.2% to the quintuple-zeta results for all studied molecules. In comparison to the other molecules, W2 converges slower with the basis set size. Angeli et al. reported that in order to reproduce the experimental values of vibrational frequencies and bond length, a third-order multi-reference calculation with a very large ANO-RCC basis set was needed [81]. The ground state of the tungsten atom has several low-lying electronic states and their combinations to low-lying electronic states in the W2 dimer was studied in detail Ref. [82] using a quadruple-zeta ANO-RCC basis set. With this excellent basis set convergence, the quadruple-zeta basis will be used for the comparative study of atomization energies, structures and vibrational frequencies for Be/W/H molecules for various DFT methods in Section 6.

5. Be2 and W2

It is known that Be2 and W2 pose special problems to a quantum chemical treatment [81,83–89]. Calculated data on atomization energies for Be2 and W2 from all tested DFT functionals are listed in Table
S2 in the supplementary material. Large deviations arise between different methods for these two dimers. Thus, we now discuss the two outlier molecules Be2 and W2 separately, and they were not included in the data set for comparing the performance of DFT functionals in Section 6.

The beryllium dimer with only 8 electrons has an unusual nature of chemical bonding. A large number of papers aiming at understanding the bonding mechanism have appeared since the 1960s. The very first ab-initio study for Be2 is a limited configuration interaction (CI) calculation by Fraga and Ransil [83], where 16 excited electronic configurations were employed. It was found that the total energy of Be2 was greatly lowered by 1.3 eV with respect to the data obtained with a single configuration. In 1983, Harrison and Handy employed a full CI calculation including 912,464 CSFs with a large Gaussian type basis set which yielded $D_e = 0.080$ eV and $r_e = 2.514$ Å [84], which is still rather far from the experimental findings by Bondybey and English [90] in 1984, who report a Be2 bond energy of 0.10 ± 0.004 eV and a bond length of 2.45 Å. Koput [86] obtained $D_e = 0.116$ eV and $r_e = 2.444$ Å from CCSD(T) calculations with the septuple-zeta-correlation-consistent core-valence basis sets including full CI corrections and scalar relativistic effects. This agrees very well with those predicted by Meshkov et al. ($D_e = 0.116$ eV, $r_e = 2.445$ Å) [85] and with recent experimental findings ($D_e = 0.115$ eV, $r_e = 2.45$ Å) [70]. It is clear that Be2 is unsuitable to be treated with the range of functionals presented here as well as with standard CCSD(T) calculations.

The tungsten dimer also poses difficulties for quantum chemical methods, and the main reasons have been summarized by Angeli et al. [81] DFT calculations by Wu and Ma [87] indicated that at a given spin multiplicity, more than one local minimum can be found for the W dimer. The global minimum is found to be $^4\Sigma_g^+$ with a bond distance of 2.039 Å, while the triplet state is very competitive and lies only 0.131 eV higher than the ground electronic state [87]. Modified Hückel calculations to study the valence and valence-core effects of W2 have been reported by Boudreaux and Baxter [88], where the bond energy is found to be 5.166 eV at a distance of 2.01 Å. Angeli et al. [81] carried out multireference perturbation theory calculations giving a bond distance $r_e = 2.056$ Å and the harmonic frequency $\omega_e = 327$ cm$^{-1}$ in fairly good agreement with the theoretical ($r_e = 2.048$ Å) [89] and experimental estimations ($\omega_e = 336.8$ cm$^{-1}$) [91]. The experimental value of the binding energy for W2 was estimated to be around 5 ± 1 eV [71]. Fortunately, the problems for Be2 and W2 do not extend to the other molecules considered here which seem to behave much more regular.

In the present work, we do not aim at a more accurate description of Be2 and W2, but leave these molecules in the calculation set while keeping in mind that DFT results for Be2 and W2 can be far off the true values. Most DFT methods give too large values for $D_0$ of Be2 but some range-separated functionals give too small values. We note that $\omega$B97XD yields the most accurate atomization energy of 0.08 eV for Be2 by comparing with the experimental value of 0.1 eV.

For the W dimer, most methods predict a singlet $\Sigma^+$ ground state in agreement with the conjecture by Hu et al. [91] when comparing W2 to its congeners Cr2 and Mo2 which have been suggested in the $\Sigma^+$ ground state with flash photolysis [92,93]. HF failed in W2 and it predicts a septet state, with an energy 1.85 eV lower than the singlet state. It was pointed out in Ref. [94] that it is hard to provide a reasonable starting orbital for the open-shell system for HF calculations. Of course, the HF method does not include any correlation effect. The 6s and 5d subshells of W are energetically close, causing large near-degeneracy correlation effects [95]. GGAs and Meta-GGAs give atomization energies for the W dimer close to experimental values (Table 1) while RSH-(meta)GGA yield too small energies.

6. Comparison of DFT functionals

DFT is established as a cost-effective approach to solve the electronic structure problem in particular for ground-state physical properties of molecules and periodic systems. We systematically test a set of DFT functionals by a direct comparison with CCSD(T) calculations. The cc-pVQZ GTO-basis set was chosen as a reasonable compromise.

We calculated the energies for the four lowest possible spin states for all molecules and atoms to find the lowest-energy electronic configuration for each functional. Different methods do not always predict the same ground state: The $^5\Sigma_0^+$ state predicted for atomic W by B97D, B97D3, HCTH, oB97X-D and M06. According to mass-selected double-resonance ionization spectroscopy [96] this is the ground state. Other methods including CCSD(T) and HF yield $^7\Sigma_0^+$ as the lowest-energy state. The compilation of energy levels for W by Kramida and Shirai [97] indicate a ground state of $^6\Sigma_0^+$, and a competitive state of $^7\Sigma_0^+$, owing to the strong mixing between $5d^66s^2$ and $5d^56s$ configurations.

In this work, we do not force the molecules or atoms to assume the experimental ground state configuration but rather report the structures, energetics, and vibrational frequencies from the configuration with the lowest energy for each method.

The molecules listed in Table S3 in the supplementary material were included in the test set for a systematic comparison of DFT methods. All molecules are in their optimized geometry. Reference CCSD(T) energies were determined from the CBS limit from cc-pVQZ/cc-pV5Z extrapolations shown in Table 2, and reference geometries and frequencies were taken from the results at the CCSD(T)/cc-pVQZ level.

6.1. Atomization energies

All atomization energies in Table S3 are zero-point-energy corrected. Mean absolute errors (MAE) per atom relative to the CBS limit from CCSD(T) cc-pVQZ/cc-pV5Z extrapolations are shown in Fig. 2.

The overall accuracy of the energies follows Jacob's Ladder [98], with meta-GGA being the exception. LSDA already reduces the large error of HF but only the inclusion of non-local correlation leads to smaller deviations between DFT and CCSD(T). Semi-local GGA methods

![Fig. 2. Mean absolute errors (MAE) per atom for atomization energies relative to the CBS limit from CCSD(T) cc-pVQZ/cc-pV5Z extrapolations for Be/H/W molecules listed in Table S3.](image-url)
give a MAE ranging from 0.07 eV (B97D) to 0.12 eV (BP86) per atom. HSEH1PBE performs slightly worse than the other functionals in rung 4. Several of the functionals employing HF exchange for long-range corrections do not give improved atomization energies for the covalently bound molecules.

All functionals in rung 4 and three functionals (B97D, B97D3 and BLYP) in rung 2 perform superior, with MAs smaller than 0.1 eV/atom. It is of interest to note that B97D3 which employs a newer version of the dispersion parameters D3(BJ) has a 41% larger MAE than B97D. The maximal error for most functionals occurs for Be4 or WBe.

6.2. Accuracy of ground state geometries

To assess the quality of structural data from the functionals, we analyze MAs for the bond lengths. The definitions of MAE for a single molecule can be written as:

\[
\text{MAE} = \frac{1}{n} \sum_{i=1}^{n} \left| R_i - R_i^{\text{ref}} \right|
\]

where \(R_i\) and \(R_i^{\text{ref}}\) are bond lengths calculated by DFT (or HF) and CCSD(T), respectively, and \(n\) is the number of bonds in the molecule. From Eq. (4), we calculated \(\text{MAE}\) by averaging over all studied molecules. We do not include bond angles in the same way because they as well as dihedral angles do not play a significant role for our molecules with four or less atoms.

The \(\langle \text{MAE} \rangle\) values are shown in Fig. 3. The most accurate functional is M11 with an MAE of 0.007 Å, closely followed by \(\omega\)B97XD and HSEH1PBE.

From Figs. 2 and 3 one sees that range-separated functionals M11, \(\omega\)B97XD and HSEH1PBE predict both atomization energies and bond lengths for the studied molecules rather well. Some methods that are accurate for predicting atomization energies are not effective for bond lengths or vice versa. For instance, CAM-B3LYP gives a relatively large error in bond lengths but is among the best performing methods for atomization energies. Conversely, the Meta-GGA functional TPSSTPSS is good for geometries but inaccurate for energies.

6.3. Harmonic vibrational frequencies

We define mean absolute relative errors (MARE) in frequencies as:

\[
\text{MARE} = \frac{1}{n} \sum_{k=1}^{n} \left| f_k - f_k^{\text{ref}} \right| / f_k^{\text{ref}}
\]

where \(f_k\) and \(f_k^{\text{ref}}\) are the unscaled vibrational frequencies in the harmonic approximation on the DFT (or HF) and CCSD(T) PES, and \(n\) is 3N-5 for linear molecules and 3N-6 for all others. MARE from Eqs. (5) are averaged over the studied molecules and mean \(\langle \text{MARE} \rangle\) values for all complexes are shown in Fig. 4.

HSEH1PBE in rung 4 proves to be the best approach for harmonic frequencies. All DFT functionals above rung 1 give errors within 5% \(\omega\)B97XD has a larger error than the other functionals in rung 4, since it predicts a too small frequency for the symmetric bending mode BeH-Be for Be3H. One can also see that the addition of long-range corrections shows very different effects for various functionals. It increases the error by 36% from B3LYP to CAM-B3LYP, in contrast decreases the error by 23% from M06 to M11.

6.4. Summary of the performance of DFT functionals

Average errors for all methods (functionals) of a specific type are shown in Fig. 5 for energies, bond lengths and frequencies.

In case of atomization energies (Fig. 5(a)), hybrid methods on average show a better performance for the studied Be/H/W molecules. The meta-GGAs are inferior to the GGAs.

When it comes to bond lengths (Fig. 5(b)), the tested GGAs and RSH-(meta)GGAs show the best performance with an error smaller than 0.01 Å on average. The meta-GGAs and H-(meta)GGAs give bigger errors of 0.015 Å.

For harmonic vibrational frequencies (Fig. 5(c)), the errors decrease steadily from the LSDA to hybrid (meta-)GGAs. The inclusion of range-
separated corrections having negative effects slightly disturb the descending trend.

In summary, hybrid (meta-)GGAs on average perform the best in predicting energetics and frequencies, GGAs are superior to the others for the bond lengths, and RSH-(meta)GGAs show a general good performance for all three properties.

7. Conclusions

CCSD(T) energies of the considered molecules were extrapolated to the CBS limit using different types of basis sets. We found that the CBS atomization energies at the 3/4 and 4/5 extrapolation with the cc-pVnZ basis do not differ much. cc-pVQZ/cc-pV5Z extrapolations to the CBS limit generally agree well with the energies at the CCSD(T)/cc-pVQZ level and available experimental results. All-electron calculations were performed for several W containing molecules to understand the effects of the core electron correlation energies. We note that an approximate treatment of the inner electrons can cause a difference of 3-5% for W hydrides and W₂.

Atomization energies, bond lengths and vibrational frequencies were obtained for various DFT functionals from rungs 1–4 of Jacob's ladder and compared to Hartree–Fock and CCSD(T) results for small neutral Be/H/W molecules. This combination covers both main group chemistry and transition metals and is relevant for material chemistry in the ITER thermonuclear reactor. The inclusion of range-separated corrections has positive or negative effects depending on the DFT functionals and the properties of interest. For atomization energies, oB97XD shows the best performance, followed by B97D, M06, B3LYP and M11. In the case of reproducing bond lengths, M11, oB97XD and HSEH1PBE are the best performers of this comparison. When it comes to vibrational frequencies, the ranking sequence is HSEH1PBE, B3LYP and M11. In summary, the range-separated hybrid mega-GGA functional M11 stands out with accurate results for all three properties for the studied Be/H/W molecules listed in Table S3.

Analyses of average errors for all methods of a specific type yielded, on average, a better performance of hybrid (meta)GGAs in predicting energetics, and frequencies than the functionals in the lower rungs and GGAs being superior to the others for bond lengths. The results presented here span a variety of unusual molecules that are normally not considered in the training or evaluation procedures for density functionals. They depend, of course, also on the accuracy of the reference CCSD(T) method which is hard to evaluate at this point and may lead to necessary reevaluations for some molecules should it turn out to be inadequate in theoretical or experimental studies for specific molecules. However, DFT methods are heavily used to predict many properties of plasma-facing materials and are also used for force field generation for dynamical simulations in our and other groups. Therefore, we find it necessary to gain insight into their predictive power and their shortcomings for these specific element combinations.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work has been funded by the Austrian Science Fund (FWF) project under Grant No. P28979-N27, by the Kommission zur Koordination der Kernfusionsforschung der ÖAW (KKKÖ) Matching Grant, and by the Tiroler Science Fund (TWF) Project UNI-0404/2308. It has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom research and training program 2014-2018 and 2019-2020 under grant agreement No 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission. The computational results presented here have been achieved using the Vienna Scientific Cluster (VSC-3) and the infrastructure LEO of the University of Innsbruck.

References


Beryllium, tungsten and their alloys Be$_2$W and Be$_{12}$W: Surface defect energetics from density functional theory calculations

L. Chen$^a$, I. Sukuba$^{a,b}$, M. Probst$^b$, A. Kaiser$^{b,*}$

$^a$ Universität Innsbruck, Institut für Ionenphysik und Angewandte Physik, Innsbruck 6020, Austria
$^b$ Department of Nuclear Physics and Biophysics, Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava SK-84248, Slovakia

A B S T R A C T

The binding, adsorption and removal energies of atoms on the surfaces of Be with the major orientations (1120) and (1121), and of W and Be/W alloys with (110) and (111) orientations have been obtained from density functional theory calculations. An analysis of adatom formation on viable surfaces and of adsorption and step formation processes is provided as well. Possible surface changes at conditions occurring during plasma-wall interactions are discussed. The large differences of the energetic parameters for various surface orientations and terminations mean that atomistic modeling is indeed essential for understanding the behavior of PFMs in a fusion environment.

1. Introduction

Beryllium and tungsten surfaces play an important role in nuclear fusion devices like the Joint European Torus (JET) tokamak [1,2] or other medium size tokamaks and will be utilized in the International Thermonuclear Experimental Reactor (ITER) [3]. Moreover, tungsten will possibly be used in the Demonstration Fusion Power Reactor (DEMO) [4]. They serve as plasma-facing materials (PFMs) at the inner walls of the vacuum vessel with beryllium at the first wall and tungsten at the divertor. Their interaction with the fusion plasma is one key interest in fusion research. It is necessary to map out the plasma-wall interactions (PWIs) under different conditions defined by the heat and particle loads and by subsequent processes like sputtering, adsorption and deposition that all influence both the plasma and the PFMs. These processes can, for example, result in the formation of alloys and the formation of Be$_2$W and Be$_{12}$W, in particular, has been observed in deposition experiments [5-8]. In situ alloy formation and other processes like outgassing, W fuzz formation, etc. would lead to degradation of the PFMs if they happen to a larger degree [9]. Several experimental studies on this topic have been performed on PWIs in tokamaks or in ion beam and linear plasma devices [10-12]. Computational modeling at various level of theory was used to support and explain experimental results, and to extrapolate towards long-term PFM stability. On the most fundamental level, the modeling starts with atomistic quantum chemical calculations with a strong emphasis on density functional theory (DFT) [13-20] and then proceeds to atomistic molecular dynamics (MD) of time-dependent processes [21,22]. Up at the multiscale ladder it further proceeds to differential equation solvers like ERO [23] or kinetic Monte Carlo [24] and to other higher-level treatments that incorporate realistic experimental geometries and plasma environments as well as long time scales.

 Atomistic surface data, such as surface binding energies $E_{\text{BB}}$ or adsorption energies can deepen our understanding of material migration, especially for sputtering and alloy formation at high energy plasma loads. Of course, in a technical device no monocrystalline surfaces are used. However, surfaces of materials like in the tiles that will actually be used consist of microcrystals and in order to fully understand their macroscopic performance, model systems have first to be studied.

In a former work, density functional theory calculations have been used to obtain $E_{\text{BB}}$ for Be(0001), W (001), Be$_2$W(001) and Be$_{12}$W(001) surfaces with various terminations in Ref. [20]. In the present work, we calculate $E_{\text{BB}}$ for several surfaces of Be, W and Be/W alloys with major orientations given by the Miller indices (110) and (111). We augment the $E_{\text{BB}}$ data with the formation energies for a single adatom on various surfaces and with some step formation energies. This data set can provide benchmark information for atomistic force fields and thus contribute to the understanding of the plasma-surface equilibrium conditions [25].

2. Details of surface models and calculations

The stoichiometric systems were constructed as slabs with surface terminations in the (110) or (111) planes. Due to the use of plane waves
as basis sets, it is of advantage to stay in a system with 3D periodicity, and therefore a slab geometry is used for our surface studies. The thickness of the vacuum, or specifically the length of vacuum along the direction perpendicular to the slab surfaces was chosen to be 8 Å. This is large enough to ensure that opposing surfaces do not disturb each other. The supercell dimensions, the supercell vectors \( \overrightarrow{a}, \overrightarrow{b}, \overrightarrow{c} \), and the number of atoms of the surface system are described in Table 1. The surfaces are oriented parallel to the plane spanned by \( \overrightarrow{b}, \overrightarrow{c} \). The detailed information of the super cell and atomic positions for these surfaces are reported in the Supplementary Material. The crystal structure of Be\(_2\)W is hexagonal (space group P6\(_3\)/mmc) [26] and Be\(_2\)W is isostructural to Be\(_2\)Mo (space group I\_4/mmm) [27].

Kohn-Sham DFT was used to predict the ground state energy and all geometries were locally optimized [28,29]. Note that local minima were not confirmed by frequency calculations and some of the structures might be transition states or higher order saddle points although that is unlikely. A Plane-Waves (PW) basis set with a kinetic cutoff energy of 550 eV was employed. The Projector Augmented Wave (PAW) method and the Perdew-Burke-\v{E}rnzerhof (PBE) exchange-correlation functional [30] describe the core and valence electrons. Standard PAW potentials for W (with six valence electrons) and Be (with two valence electrons) were taken from the VASP library. In most of the calculations, a Gamma-centered k-point mesh of \( 5 \times 5 \times 5 \) was employed if not otherwise indicated below. The supercell dimensions in \( \overrightarrow{a}, \overrightarrow{b}, \overrightarrow{c} \) directions are of comparable size for the calculation of surface defects (Table 1). In the geometry optimizations the two lower layers of atoms were fixed while all other atoms were allowed to relax and a convergence criterion of \( 10^{-6} \) eV/Å on the forces was employed. Convergence tests for the cutoff energy and k-point grids have been performed. The lattice parameters for pure materials and alloys optimized in Ref. [20] were used. All calculations were carried out using the Vienna Ab Initio Simulation Package (VASP) [31,32].

### 3. Surface defect energetics

Surface energetics can be very different from bulk due to the broken periodicity and the lower coordination of surface atoms. For alloys, an additional complication comes from the fact that a number of surface terminations exists for every cutting direction, each with a different stability. The usual definitions of surface energy or step-formation energy are almost meaningless for periodic surface slab calculations of alloys. First, most slabs consist of two different surfaces, and second, a cohesive energy per atom type cannot easily be defined. However, the surface binding energy is a physically meaningful quantity that can be perfectly defined for surfaces of alloys with the common definition:

\[
E_{\text{SBE}} = E_{\text{atom}} + E_{\text{vac}} - E_{\text{S}}
\]

where \( E_{\text{vac}} \) is the total energy of the optimized surface with a single surface vacancy, \( E_{\text{S}} \) is the total energy of the clean surface, and \( E_{\text{atom}} \) is the atomic energy of the removed atom. The electronic configurations of the reference atoms are the singlet \( 1\text{S} \) (\( 1s^22s^2 \)) state for Be, and the septet \( \text{5S}_9 \) (\( \text{[Xe]}4f^{14}5d^66s^1 \)) for W. The removal of a surface atom in the top layer also changes the stoichiometry of the slab.

The vacancy formation energy \( E_{\text{vac}} \) is defined with reference to the average energy of a bulk atom \( E_{\text{bulk}} \):

\[
E_{\text{vac}} = E_{\text{atom}} + E_{\text{vac}} - E_{\text{S}}
\]

For each possible surface termination, non-equivalent single atom vacancy sites exist and were labeled with Greek letters (\( \alpha, \beta, \gamma, \delta, \epsilon \)). We insert numbered superscripts (“1” for Be\(_2\)W(110), “2” for Be\(_2\)W(111), “3” for Be\(_2\)W(110) and “4” for Be\(_12\)W(111)) after the Greek letters to facilitate the distinction between different surface structures, i.e., Be(\( \alpha^1 \)) is a Be vacancy in Be\(_2\)W(110). The vacancies are labeled in the top-most layer of a surface. In case of Be(\( \alpha^1 \)), the top-most layer is Slab A as shown in Fig. 1. Thus, the Greek letters denote positions with different neighborhoods but their labeling sequence was chosen without physically motivated background. In Be\(_2\)W(110) surface systems, two possible terminations denoted as Slab A, B have three non-equivalent vacancies of Be(\( \alpha^1 \), \( \beta^1 \), \( \gamma^1 \)) and a single vacancy of W(\( \delta^1 \)), as shown in Fig. 1. The super cell of Be\(_2\)W(110) surface is shown in Fig. 2, which illustrates six different terminations (Slab A-F) with six Be (\( \alpha^1, \beta^1, \gamma^1 \)) and one W(\( \delta^1 \)) vacancy. Surface systems with non-equivalent vacancies for Be\(_2\)W(111) and Be\(_12\)W(111) are individually shown in the Supplementary Material in Figs. S1 and S2.

The calculated surface binding energies and vacancy formation energies are summarized in Table 2 together with the number of nearest neighbors (NNs) of the vacancies and the distances between vacancies and their nearest neighbors. The beryllium \( E_{\text{SBE}} \) values are in a range of \( 4.45 \pm 1.18 \) eV for all structures. The difference of 2.36 eV between the largest and smallest one means that models that do not account for the atomistic surface structure will be substantially inaccurate. It is noticeable that Be(\( \alpha^3 \)) for the (110) surface and Be(\( \lambda^2 \)) for the (111) surface of Be\(_2\)W are bound less strongly than all other Be vacancies with an \( E_{\text{SBE}} \) of 3.3 eV. Be

![Fig. 1.](image-url)
atoms are strongest bound in Be\textsubscript{2}W(001) surfaces with an ESBE\textsubscript{5.63} eV. The differences in ESBE can be attributed to the variety in the atomic neighborhood of the vacancies. For instance, both Be(\(\gamma_{3}\)) and Be(\(\delta_{3}\)) have six same Be nearest neighbors, while the latter holds one more W neighbor than the former, which means that one more Be-W bond has to be broken to remove Be(\(\delta_{3}\)) from the top surface than to remove Be(\(\gamma_{3}\)). Correspondingly, ESBE for Be(\(\delta_{3}\)) is 0.54 eV higher than that for Be(\(\gamma_{3}\)). Identically, Be(\(\gamma_{4}\)) and Be(\(\beta_{4}\)) have the same W nearest neighbors, \(\text{Fig. 2.}\)

**Table 2**

Surface binding energies ESBE and vacancy formation energies \(E_{\text{vac}}\) number of nearest neighbors (NNs) and their distances for various surface terminations of Be, W, and two of their alloys.

<table>
<thead>
<tr>
<th>(\text{Be}<em>{x}\text{W}</em>{y})</th>
<th>vacancy</th>
<th>Number of NNs</th>
<th>(E_{\text{vac}}) (eV)</th>
<th>(E_{\text{Be}}) (eV)</th>
<th>(E_{\text{W}}) (eV)</th>
<th>Distances [nm] (number of neighbors in parentheses)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Be(0001)\textsuperscript{a}</td>
<td>Be</td>
<td>9</td>
<td>-</td>
<td>1.401</td>
<td>5.128</td>
<td>0.221(3), 0.228(6)</td>
</tr>
<tr>
<td>Be(1120)</td>
<td>Be</td>
<td>7</td>
<td>-</td>
<td>0.644</td>
<td>4.371</td>
<td>0.228(3)</td>
</tr>
<tr>
<td>Be(1011)</td>
<td>Be</td>
<td>6</td>
<td>-</td>
<td>0.281</td>
<td>4.008</td>
<td>0.228(3)</td>
</tr>
<tr>
<td>W(001)\textsuperscript{b}</td>
<td>W</td>
<td>-</td>
<td>4</td>
<td>0.320</td>
<td>8.805</td>
<td>0.228(3)</td>
</tr>
<tr>
<td>W(110)</td>
<td>W</td>
<td>-</td>
<td>6</td>
<td>1.787</td>
<td>10.272</td>
<td>0.228(3)</td>
</tr>
<tr>
<td>W(111)</td>
<td>W</td>
<td>-</td>
<td>4</td>
<td>0.244</td>
<td>8.729</td>
<td>0.228(3)</td>
</tr>
<tr>
<td>Be\textsubscript{2}W(001)</td>
<td>Be</td>
<td>5</td>
<td>3</td>
<td>1.900</td>
<td>5.627</td>
<td>0.218(2), 0.225(1), 0.227(2)</td>
</tr>
<tr>
<td>W</td>
<td>6</td>
<td>3</td>
<td>0.274(1), 0.262(2)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Be\textsubscript{2}W(110)</td>
<td>Slab A Be((\alpha_{1}))</td>
<td>4</td>
<td>2</td>
<td>0.948</td>
<td>4.675</td>
<td>0.225(2), 0.227(2)</td>
</tr>
<tr>
<td>Slab B Be((\beta_{1}))</td>
<td>4</td>
<td>4</td>
<td>1.208</td>
<td>4.935</td>
<td>0.228(3)</td>
<td></td>
</tr>
<tr>
<td>Slab B W((\delta_{1}))</td>
<td>7</td>
<td>3</td>
<td>0.756</td>
<td>9.244</td>
<td>0.228(3)</td>
<td></td>
</tr>
<tr>
<td>Slab B W((\epsilon_{3}))</td>
<td>6</td>
<td>3</td>
<td>0.509</td>
<td>8.994</td>
<td>0.228(3)</td>
<td></td>
</tr>
<tr>
<td>Be\textsubscript{12}W(001)</td>
<td>Slab A Be((\alpha_{2}))</td>
<td>7</td>
<td>1</td>
<td>1.013</td>
<td>4.740</td>
<td>0.218(3), 0.225(1), 0.227(2)</td>
</tr>
<tr>
<td>Slab B Be((\beta_{2}))</td>
<td>6</td>
<td>1</td>
<td>0.274</td>
<td>4.977</td>
<td>0.228(3)</td>
<td></td>
</tr>
<tr>
<td>Slab F W((\epsilon_{4}))</td>
<td>12</td>
<td>-</td>
<td>-</td>
<td>0.274</td>
<td>8.213</td>
<td></td>
</tr>
<tr>
<td>Be\textsubscript{12}W(110)</td>
<td>Slab A Be((\alpha_{3}))</td>
<td>7</td>
<td>-</td>
<td>-</td>
<td>0.402</td>
<td>3.326</td>
</tr>
<tr>
<td>Slab B Be((\beta_{3}))</td>
<td>5</td>
<td>1</td>
<td>0.725</td>
<td>4.452</td>
<td>0.228(3), 0.227(2)</td>
<td></td>
</tr>
<tr>
<td>Slab C Be((\gamma_{3}))</td>
<td>6</td>
<td>1</td>
<td>0.708</td>
<td>4.335</td>
<td>0.228(3), 0.227(2)</td>
<td></td>
</tr>
<tr>
<td>Slab D Be((\delta_{3}))</td>
<td>7</td>
<td>1</td>
<td>0.815</td>
<td>4.947</td>
<td>0.228(3), 0.227(2)</td>
<td></td>
</tr>
<tr>
<td>Slab E Be((\epsilon_{3}))</td>
<td>5</td>
<td>1</td>
<td>0.578</td>
<td>4.306</td>
<td>0.228(3), 0.227(2)</td>
<td></td>
</tr>
<tr>
<td>Slab F W((\epsilon_{4}))</td>
<td>10</td>
<td>2</td>
<td>0.332</td>
<td>9.309</td>
<td>0.228(3), 0.227(2)</td>
<td></td>
</tr>
<tr>
<td>Slab E Be((\gamma_{3}))</td>
<td>6</td>
<td>2</td>
<td>0.274</td>
<td>8.213</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Slab D W((\epsilon_{4}))</td>
<td>10</td>
<td>-</td>
<td>-</td>
<td>0.274</td>
<td>8.213</td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{a} ESBE of Be (0001), W (001) reported in Ref. [20].

\textsuperscript{b} ESBE of the more stable Be or W atom on Be\textsubscript{2}W(001) and Be\textsubscript{12}W(001) surfaces reported in Ref. [20].
while Be(β4) has one more Be neighbor resulting in the higher \( \text{ESBE} \) of Be(β4) in comparison with Be(γ4). The distinction of \( \text{ESBE} \) between Be(γ4) and Be(β4) can also be reasonably explained by the difference of their Be neighbors.

When it comes to the tungsten \( \text{ESBE} \), results are between the minimal value of 6.81 eV for Be12W(001) and the maximal one of 10.27 eV for W(110). The surface atom of W in (110) direction is surrounded by six other W atoms, while in Be12W(001) surfaces, it is surrounded by Be nearest neighbors only. Previous experiments estimated the binding energy of the W dimer to be \( 5 \pm 1 \) eV \([33,34]\), which strongly exceeds the value of the Be-W dimer (1.82 eV) predicted by DFT theory \([35]\). This qualitatively coincides with the trends of the surface binding energies. Although one would assume that a metallic environment can be rather different from small molecular clusters, we observed that even slight differences can be tracked back to the arrangement of atoms on the surface. For example, although W(β3) and W(γ5) have very similar W neighbors, \( \text{ESBE} \) of W(β3) is 0.25 eV higher than that of W(γ5), because W(β3) is surrounded by one more Be atom. Identical neighborhoods such as the neighborhood of the surface atom of W in (001) direction and of W in (111) direction also lead to similar \( \text{ESBE} \) values.

In summary, W atoms of Be12W surfaces are less strongly bound than Be-W, with the apparent reason of Be-W having higher W content. However, Be atoms in Be12W can be either stronger or weaker bound than in BeW.

The values of \( \text{ESBE} \) and the ones of \( \text{Evac} \) follow similar trends; higher vacancy formation energies correspond to stronger surface binding energies. Despite values for \( \text{Evac} \) for different alloys being hardly comparable, they are reported because they can serve as reference energies that are independent of the atomic spin-state for the construction of force fields.

4. Adatoms

Adatoms on metal surfaces (Pt \([36,37]\), Au \([38]\), Al \([39]\), Si \([40]\), W \([41–43]\)) have attracted much interest from both experiments and calculations because of their role as nucleation centers for thin film growth on the surface. In the present work, initial configurations were constructed by the addition of a single adatom on the surface terminations defined above and then the whole system was relaxed. The adatoms were labeled by \( \xi, \eta, \theta, \sigma, \tau, \chi, \kappa \) with a numbered superscript to indicate different alloy surfaces in the same way as for vacancies. The adatom was initially placed at existing crystal periodic positions. In Fig. 3, we show four possible positions for adatoms of Be(\( \xi \)), (\( \eta \)), (\( \theta \)), and W(\( \tau \)) on the two terminations of Be\( \chi \)W(110). The six topmost layers of the Be2W(110) super cell are shown in Fig. 4 and include seven possible positions for adatoms labeled with Be(\( \xi \)), Be(\( \eta \)), and W(\( \tau \)). The labeling of adatoms follows the same systematics as the labeling of vacancies for direct comparability. In addition, some possibilities of adatoms on Be, W, and Be/W alloy (001) surfaces were also studied as reference data. Geometries for other surfaces are illustrated in the Supplementary Material Figs. S3 and S4.

The adatom formation energy \( \text{E_{ad}} \) and binding energy \( \text{E_{bad}} \) are defined as

\[
\text{E}_{\text{ad}} = \text{E}_{\text{ad,calc}} - \text{E}_{\text{bulk}} - \text{E}_{\text{ad}}
\]

\[
\text{E}_{\text{bad}} = \text{E}_{\text{atom}} + \text{E}_{\text{ad}} - \text{E}_{\text{ad,calc}}
\]

where \( \text{E}_{\text{ad,calc}} \) is the energy of the surface with the adatom. From these definitions, \( \text{E}_{\text{ad}} \) is the energy needed to remove an atom from the bulk and then add it on top of the surface whereas \( \text{E}_{\text{bad}} \) denotes the energy required to remove the adatom from the surface to the vacuum \([44]\).

We also studied a single Be adatom on a pure W surface and vice versa. For example, in the fusion environment, a Be atom coming from the first wall due to plasma sputtering may be transported and then be deposited and adsorbed on the divertor W surface. We only calculated \( \text{E}_{\text{bad}} \) for this impurity adatom.

The values for \( \text{E}_{\text{ad}} \) and \( \text{E}_{\text{bad}} \) for different structures are summarized in Table 3, where the \( \text{ESBE} \) values from Table 2 are also included. As mentioned before, the adatom is placed in the corresponding site of the surface vacancy for comparability of the values of \( \text{ESBE} \) and \( \text{E}_{\text{bad}} \). The adatom binding energies are smaller than or roughly equal to the surface binding energy for comparable structures, indicating that adatoms are easier sputtered than surface atoms. This is not unexpected but it is important since it relates to the equilibrium governing the growth and disappearance of layers on surfaces. To a first explanation, the number of nearest neighbors for adatoms is less than that for surface atoms. In analogy to the \( \text{ESBE} \) values, W adatoms on surfaces of the Be12W structure are more weakly bound than in structures with more W content while Be adatoms are in most cases stabilized by alloy formation with tungsten. Clearly, the values of \( \text{E}_{\text{ad}} \) and the ones of \( \text{E}_{\text{bad}} \) follow opposite trends. In particular, lower adatom formation energies come from higher adatom binding energies. The adatom binding energies can be very different for different surface cuts. The adatom binding energy of W(\( \tau \)) on Be12W(110) surface is 2 eV higher than that for W(\( \tau \)) on Be12W(111) surface. On the contrary, for W adatoms on the Be12W surface, the adatom binding energies for W(\( \tau \)) and W(\( \tau \)) are within only 0.3 eV difference. Interestingly, we find a negative value for \( \text{E}_{\text{ad}} \) and the highest value for \( \text{E}_{\text{bad}} \) for the W(001) surface, probably due to a very strong relaxation of the structure with the adatom. The bond lengths between the W adatom and its nearest neighbors are 2.54 Å on W(001) and 2.44 Å on W(111), in good agreement with the values of 2.45 Å and 2.59 Å obtained by Xu using fourth moment tight-binding theory \([45]\).

Another interesting observation is that \( \text{E}_{\text{ad}} \) of W adatoms on the Be surfaces is much larger than that of Be adatoms on the W surfaces, indicating a larger stability of W impurities deposited on Be surfaces. This weak binding of Be adatom is qualitatively consistent with findings from beryllium seeding experiments by Doerner \([5]\), where 0.1% of beryllium was accumulated on W surface. The binding energy of the W adatom on the Be(0001) surface is 3.8 eV, agreeing well with the DFT calculations (4.2 eV) by Allouche with a similar method \([46]\). The binding energy of the Be adatom on the W(001) surface is 1.38 eV, which is smaller than the data in Ref. \([47]\) due to the relaxation of the surface in our calculation.

Again, one sees that the dependence of the adatom binding energies and of the surface binding energies on the specific surface cuts and surface terminations makes an atomistic modeling essential.
5. Step formation

Steps are often observed first-order surface defects and we chose to investigate some simple step models here. In particular, we added a monolayer and removed a part of the topmost atoms of the surface slab to generate two steps per surface. The step system was relaxed freely to determine the minimum total energy $E_{\text{step}}$. The step formation energy per length $G_{\text{step}}$ is defined for pure materials as:

$$G_{\text{step}} = \frac{E_{\text{step}} - N_{\text{Be}} E_{\text{Be}}^{\text{bulk}} - N_{\text{W}} E_{\text{W}}^{\text{bulk}} - E_S}{2L}$$  \hspace{1cm} (5)$$

where $N_{\text{Be}}$ and $N_{\text{W}}$ are the numbers of Be and W atoms respectively added to the surface to form the steps, and $L$ is the length of the step. $G_{\text{step}}$ represents the energy needed to remove several atoms from the bulk and then add them collectively on the top of the surface for the formation of a specific double-step.

The binding energy of a ledge atom is defined in analogy to the surface binding energy and the adatom binding energy and is, in contrast to the step formation energy, well defined and comparable for alloys. The binding energy of a ledge atom $E_{\text{bled}}$ in a step is calculated as:

$$E_{\text{bled}} = E_{\text{step}+V} + E_{\text{atom}} - E_{\text{field}}$$  \hspace{1cm} (6)$$

Here $E_{\text{step}+V}$ denotes the energy for the step system with a single vacancy in the step. $E_{\text{field}}$ demonstrates the effort to remove the ledge atom from the step to vacuum.

Twice the cell sizes given in Table 1 were used for the computations of step formations on the Be$_2$W surface and a k-point mesh of $1 \times 1 \times 1$ was employed in these cases.

Two non-equivalent steps A, B along [110] direction were created on the Be$_2$W (110) and Be$_{12}$W (110) surface systems, as shown in Figs. 5 and 6.

The step formation energy $G_{\text{step}}$ as well as the ledge atom binding energy $E_{\text{bled}}$ are given in Table 4. Our results reflect that step formation on the W(110) surface should proceed more easily along the [111] direction (174 meV/Å) than on the other two ones (220 and 223 meV/Å). A comparison with previous estimates by Xu gives a $G_{\text{step}}$ for the [111] step of 136 meV/Å [25]. The difference between the [111] direction and the other ones can be explained by the different distribution of atoms on the steps; one bond less is broken for the generation of the [111] step than for the other two steps. The values of $E_{\text{bled}}$ for those three steps show, as expected, the reversed trend; the ledge atom on the [111] step has the highest binding energy, as more bonds to neighbors have to be broken for the removal of the ledge atom from the step.

Two steps comprised of pure Be atoms on (110) surfaces of the Be$_2$W and Be$_{12}$W structure show negative step formation energies, in contrast...
to steps consisting of a mixture of Be and W atoms on these structures which have high values of $G_{\text{step}}$ (up to 250 meV/Å). The steps considered represent only a small fraction of the possible steps in the alloy systems.

Notably, the ledge atom binding energy is between the adatom binding energy and the surface binding energy for comparable structures, proving the expected stability order.

### 6. Discussion

As mentioned above, the values of $E_{\text{ad}}$, $E_{\text{tot}}$, or $E_{\text{surf}}$ correlate strongly with the types of the nearest neighbors of the removed atom. Without distinguishing the type of the binding energy but taking into account the type of the removed atom (Be or W), and without considering distances, we plot these energies from different surface systems against the number of the nearest neighbors in Fig. 7.

W neighbors significantly influence energies to remove Be atoms. In the condition of no W neighbors, Be atoms of Be$_2$W surface systems are less strongly bound than for pure Be (see Fig. 7a dash circled area 1). The reason is that Be atoms are more close-packed in pure Be than in Be$_2$W, leading to shorter distances between Be atoms. Conversely, Be atoms of the Be$_2$W surface systems are in most cases more stable than the ones of Be$_2$W. For example the energy for removing the Be atom of Be$_2$W is 0.9 eV higher than for the atom of Be$_2$W, although both of them hold four Be and two W nearest neighbors (see Fig. 7a position 2).

The energies to remove W atoms (see Fig. 7b) range from 4.77 eV for the atom with four W nearest neighbors in Be$_2$W to 10.27 eV for the atom with six W nearest neighbors in pure W. Similarly, the number of W neighbors exerts influences on these energies. Generally, the energies go up with the W content in the material: W atoms of pure W surface systems are more stable than in Be$_2$W, and those are more stable than in Be$_2$W.

Although we can see a correlation between stability and the number and types of neighbor atoms, the situation can be more complex in metals and a detailed view on the electronic structure may shed light on the stability criteria. Here, we compare electronic local densities of states (DOS) on the 5d projected orbitals of the various chemically different tungsten atoms, where the Fermi level is taken as the reference of energy. The DOS contributions of a W atom in the pure W(110) surface attributed to 5d orbitals and of a Be atom in pure Be(1120) attributed to 2p orbitals are shown in Fig. 8a.

In Fig. 8b, W(c$^3$) denotes the tungsten surface atom corresponding to the c$^3$ vacancy in the Be$_2$W(110) surface of the alloy (defined in Fig. 2), and it has twelve close Be neighbors. W(c$^3$) is the tungsten adatom on the same surface (defined in Fig. 4), which has four less Be neighbors than W(c$^3$). The highest DOS peak for the pure tungsten surface is at $\varepsilon = -3.01$ eV, while in the Be$_2$W alloy, this peak is shifted to $\varepsilon = -1.15$ eV for the adatom W(c$^3$) which indicates W-Be orbital hybridization. Due to a larger contribution of Be atoms, the peak for surface atom W(c$^3$) is even more shifted to $\varepsilon = -0.61$ eV; on the contrary, the surface binding energy of W(c$^3$) is 0.16 eV larger than the adatom binding energy of W(c$^3$). It indicates that the number of neighbors is at least as or even more important than the changes in the electronic structure upon alloy formation in this case.

In Fig. 8c, W(b$^3$) is the tungsten surface atom corresponding to the b$^3$ vacancy in the Be$_2$W(110) surface (defined in Fig. 1), and W(c$^3$) is the tungsten adatom on the same surface (defined in Fig. 3). W(b$^3$) has four more neighbors (two W and two Be) than W(c$^3$). In the Be$_2$W alloy, the highest DOS peak of W is shifted to $\varepsilon = -1.7$ eV for the surface atom W(b$^3$) and to $\varepsilon = -1.4$ eV for the adatom W(c$^3$), respectively. The Be neighbors cause a right shift, while W neighbors lead to a left shift. In this case, the binding energy of the surface atom W(b$^3$) is almost 3 eV larger than for the adatom W(c$^3$) which can be attributed mainly to two more tungsten neighbors and can hardly be inferred from the subtle changes in the projected DOS. However, it seems quite clear by comparison of Fig. 8b and c that W atoms in the W-rich Be$_2$W alloy are stronger bound.
than in Be$_{12}$W due to more states at lower energies in the DOS.

7. Conclusions

We calculated surface binding energies $E_{\text{SRH}}$ and vacancy formation energies $E_{\text{vac}}$ for single atom vacancies on the surfaces of Be with Miller indices (1120) and (1121), and W and two Be/W alloys with Miller indices (110) and (111). The values of $E_{\text{SRH}}$ are $4.45 \pm 1.18$ eV for beryllium and $8.54 \pm 1.73$ eV for tungsten, depending strongly and regularly on their neighborhood. We also investigated the energetics of addition of a single adatom at the existing crystal periodic positions on various surface terminations in order to obtain adatom binding energies $E_{\text{bad}}$ which are an essential parameter to evaluate deposition and adsorption processes. A correspondence between the values of $E_{\text{SRH}}$ and $E_{\text{bad}}$ exists: W adatoms on surfaces of the Be$_{12}$W structure are less stable than in Be$_{12}$W. Moving to higher-order defects, specific double-steps have been generated by adding several atoms on surface terminations Slab B and A (shown in Fig. 2), respectively.

Table 4

<table>
<thead>
<tr>
<th>Be/W alloy</th>
<th>Step orientations</th>
<th>$G_{\text{SRH}}$ (meV/Å)</th>
<th>Ledge atom</th>
<th>$E_{\text{bad}}$ (eV)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Be(0001)</td>
<td>[010]</td>
<td>95</td>
<td>Be</td>
<td>4.770</td>
</tr>
<tr>
<td>Be(120)</td>
<td>[0001]</td>
<td>54</td>
<td>Be</td>
<td>2.965</td>
</tr>
<tr>
<td>Be(121)</td>
<td>[1200]</td>
<td>125</td>
<td>Be</td>
<td>3.636</td>
</tr>
<tr>
<td>W(100)</td>
<td>[001]</td>
<td>10</td>
<td>W</td>
<td>8.837</td>
</tr>
<tr>
<td>W(110)</td>
<td>[001]</td>
<td>220</td>
<td>W</td>
<td>9.208</td>
</tr>
<tr>
<td>W(111)</td>
<td>[110]</td>
<td>174</td>
<td>W</td>
<td>9.465</td>
</tr>
<tr>
<td>W(0001)</td>
<td>[001]</td>
<td>–185</td>
<td>W</td>
<td>9.403</td>
</tr>
<tr>
<td>Be$_2$W(110)</td>
<td>[110]</td>
<td>–146</td>
<td>Be</td>
<td>4.449</td>
</tr>
<tr>
<td>Be$_2$W(111)</td>
<td>[110]</td>
<td>87</td>
<td>Be</td>
<td>4.564</td>
</tr>
<tr>
<td>Be$_2$W(001)</td>
<td>[110]</td>
<td>–99</td>
<td>Be</td>
<td>4.262</td>
</tr>
<tr>
<td>Be$_2$W(011)</td>
<td>[150]</td>
<td>–24</td>
<td>W</td>
<td>9.108</td>
</tr>
<tr>
<td>Be$_2$W(110)</td>
<td>[100]</td>
<td>209</td>
<td>Be</td>
<td>4.356</td>
</tr>
<tr>
<td>Be$_2$W(111)</td>
<td>[110]</td>
<td>–156</td>
<td>Be</td>
<td>4.621</td>
</tr>
<tr>
<td>Be$_2$W(111)</td>
<td>[150]</td>
<td>250</td>
<td>W</td>
<td>7.360</td>
</tr>
</tbody>
</table>

* The way to construct these steps is shown in the supplementary material in Figs. S5 and S6.

than in Be$_{12}$W due to more states at lower energies in the DOS.

7. Conclusions

We calculated surface binding energies $E_{\text{SRH}}$ and vacancy formation energies $E_{\text{vac}}$ for single atom vacancies on the surfaces of Be with Miller indices (1120) and (1121), and W and two Be/W alloys with Miller indices (110) and (111). The values of $E_{\text{SRH}}$ are $4.45 \pm 1.18$ eV for beryllium and $8.54 \pm 1.73$ eV for tungsten, depending strongly and regularly on their neighborhood. We also investigated the energetics of addition of a single adatom at the existing crystal periodic positions on various surface terminations in order to obtain adatom binding energies $E_{\text{bad}}$ which are an essential parameter to evaluate deposition and adsorption processes. A correspondence between the values of $E_{\text{SRH}}$ and $E_{\text{bad}}$ exists: W adatoms on surfaces of the Be$_{12}$W structure are less stable than in structures with more W content and Be adatoms are in most cases following the same trend. In terms of impurity adatoms, values of $E_{\text{bad}}$ indicate larger stability of W impurities deposited on Be surface. Moving to higher-order defects, specific double-steps have been generated by adding several atoms on surface terminations Slab B and A (shown in Fig. 2), respectively.

Fig. 6. Side view (ac projection) of six topmost layers in Be$_{12}$W supercell. Steps A and B along [110] direction on the Be$_{12}$W(110) surface were generated by adding several atoms on surface terminations Slab B and A (shown in Fig. 2), respectively.

Fig. 7. Energy needed to remove one Be(W) atom from Be, W, Be$_2$W and Be$_{12}$W surface systems against the number of the nearest neighbors for the removed atom.
generated on the top of the surface systems, and the ledge atom binding energy $E_{\text{bled}}$ has been calculated. The values of $E_{\text{bled}}$ correspond only weakly with $E_{\text{ads}}$, but better with $E_{\text{disf}}$ for comparable structures, in line with the observation that adatoms are easier sputtered than surface and ledge atoms. The large differences of these three energetic parameters for various surface orientations and terminations mean that atomistic modeling is indeed essential for understanding the behavior of PFMs in a fusion environment. We extended our discussion towards the DOS of W atoms in the systems of pure materials and the alloys, where the differences of the DOS can be related to their various chemical environments (Be and W neighbors). As a next step, we plan to use the data obtained in this work to generate a model to be used in dynamic simulations.

Conflict of interest

There is no conflict of interest involved in this study. All authors confirm that they are free of any conflicts of interest concerning this study.

Acknowledgments

This work has received funding from the Austrian Science Fund (FWF) project under Grant No. P289790. It has been carried out within the framework of the EUROfusion Consortium and has received funding from the Euratom Research and Training Programme 2014-2018 under Grant Agreement No. 633053. It was further supported by the Austrian Ministry of Science BMWF as part of the University Infrastructure Program of the Scientific Computing Platform LFU Innsbruck.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.nme.2018.06.021.

Fig. 8. DOS contributions of valence electrons of various atoms: a) a W atom in the W(110) surface and a Be atom in the Be(1120) surface: we used a double vertical axis (W DOS corresponding to the left vertical and Be DOS to the right); b) a surface atom W($\varepsilon$) and an adatom W($\kappa$) in the Be$_{12}$W(110) surface; c) a surface atom W ($\delta$) and an adatom W($\iota$) in the Be$_{1}$W(110) surface.

References

[21] X. Yang, A. Hassanein, Atomic scale calculations of tungsten surface binding energy


A neural network interface for DL_POLY and its application to liquid water

Ivan Sukuba, Lei Chen, Michael Probst & Alexander Kaiser

To cite this article: Ivan Sukuba, Lei Chen, Michael Probst & Alexander Kaiser (2018): A neural network interface for DL_POLY and its application to liquid water, Molecular Simulation, DOI: 10.1080/08927022.2018.1560440

To link to this article: https://doi.org/10.1080/08927022.2018.1560440

© 2018 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.

View supplementary material

Published online: 27 Dec 2018.

Submit your article to this journal

Article views: 603

View related articles

View Crossmark data
A neural network interface for DL_POLY and its application to liquid water*

Ivan Sukuba, Lei Chen, Michael Probst and Alexander Kaiser

Institute of Ion Physics and Applied Physics, University of Innsbruck, Innsbruck, Austria; Department of Nuclear Physics and Biophysics, Faculty of Mathematics, Physics and Informatics, Comenius University, Bratislava, Slovakia

ABSTRACT
After a general discussion of neural networks potential energy functions and their standing within the various approaches of representing the potential energy function of a system, we describe a new interface between the open source atomistic library aenet of Arthith and Urban and the DL_POLY 4 code. As an application example, the training of a neural network for liquid water is described and the network is used in a molecular dynamics simulation. The resulting thermodynamic properties are compared with those from a reference simulation with the same SPC/E model that has been used in the training.

1. Introduction
This chapter attempts to position neural network potential energy functions within the vast area of possible potential energy expressions and to compare them with other approaches. A suitable representation of the potential energy surface of the system under consideration is needed as input to any classical MD simulation. One can think of this representation to be the intermediate layer between the charges (which cause it) and the nuclei (which are governed by it). It can be expressed as a sum of analytical terms or it can be tabulated in some way. An early and very successful way to approximate the potential energy surface is to treat atoms connected by one up to three bonds separate from other atoms, including those in other molecules. Harmonic terms are applied for bonds and bond angles and angular functions describe dihedral energy contributions. Energy contributions arising from pairs of atoms in the same molecule but not bonded to each other or in a diatomic MD simulation. One can think of this representation to be the intermediate layer between the charges (which cause it) and the nuclei (which are governed by it). It can be expressed as a sum of analytical terms or it can be tabulated in some way. An early and very successful way to approximate the potential energy surface is to treat atoms connected by one up to three bonds separate from other atoms, including those in other molecules. Harmonic terms are applied for bonds and bond angles and angular functions describe dihedral energy contributions. Energy contributions arising from pairs of atoms in the same molecule but not bonded to each other or in a diatomic approach are the various force fields that were developed after 1970 [1], and their applications. It should be pointed out that hydrogen bonds can also be described reasonably well in that way, due to their predominantly electrostatic nature.

This approach has, however, well-known limitations that are often accepted for the sake of simplicity and computational efficiency. The main limitations are the restriction to the pair approximation (except for bond and dihedral angles) and the impossibility of breaking chemical bonds. The latter is obviously due to the harmonic terms for bound atoms. After early attempts to correctly account for bond formation and breaking within the pair approximation it turned out that this is not possible. For example, in Ref. [2] an O-H pair potential function was used that allowed H to dissociate from O (or OH). An H-H pair potential was used to account for the H-O-H angle in intact water. To accomplish this, however, this H-H pair potential included an unphysical ‘kink’ and furthermore, the auto-dissociation constant of H2O was much overestimated. Therefore, this ‘central force’ potential was not really successful for situations where the dissociation of water is important. Therefore, especially in the realm of computational materials science it is important to move beyond this stage.

There are many attempts to create potential energy functions that overcome these limitations. For example, using a ‘valence bond’ – approach [3], bonding and nonbonding energy curves can be combined to a function that incorporates bond breaking. Potential energy functions that account for bond formation or breaking are often called ‘reactive potentials’. On the other side, one way beyond the pairwise approximation is to incorporate atomic dipole polarizability. This is conceptually straightforward and can be done analytically [4] but is more often implemented with the polarisation shell model [5]. Polarisable force fields currently enjoy a renaissance, because the computer power is now available to employ them in large biophysical simulations [6].

Experience has shown that these methods, alone or combined, are often not flexible enough for the simulation of condensed matter. A certain set of analytical functions and parameters may well describe particular situations in terms of stoichiometry, structure, pressure and temperature and pairwise approximations may even include higher order contributions implicitly. However, different situations afford a
multitude of different analytical functions, if one finds a good fit at all, and as soon as it comes to mixing of force fields using combining rules, the transferability of the parameters is questionable. Therefore, other approaches prevailed. One way beyond the pairwise approximation is by tuning a pairwise interaction with a function of the coordination number of a particle. This can be called ‘pair functional’ instead of ‘pair potential’. The Sutton-Chen [7] form of the Finnis-Sinclair potential [8] which is often used for metals belongs to this class. The accuracy of this approach is still limited because there is no term modelling directed chemical bonds. To achieve this, a pairwise potential energy function is often augmented with three-body terms, typically giving a contribution only inside a sphere around the middle particle. Combining these two extensions of the pair approximation, one ends up at the stage of the so-called ‘cluster functionals’. The Tersoff-Abell potential [9,10] and other bond order potentials as the ReaxFF force field [11] belong to this class. Recently, approximate density functional calculations based on the tight-binding approximation [12] have made progress as well and it remains to be seen how they compete with other approaches.

The contributions to the potential energy beyond the pair approximation have complicated functional forms and it is by no means clear which mathematical expressions are best suited for them. On the other side, it was known since long that neural networks are about the most flexible constructions that can be used to fit arbitrary functions. This has naturally led to attempts to use neural networks in this respect, although their main application was always more towards classification and not the reproduction of exact numbers. Early attempts to use neural networks as potential energy functions worked well [13] but had no big impact because there were more efficient ways to achieve the same goal.

The present renaissance of neural network potentials for atomistic simulations can be deducted from their use in the spirit of the ‘cluster functionals’ mentioned above. A central atom interacts with its surroundings in a complicated way governed by the radial and angular density of its neighbours. The modern implementations of neural networks [14] start at this level. Each type of atom is assigned its own neural network that calculates the energy contribution from this atom. The input are the coordinates of the neighbours. In case of the ‘cluster functionals’ (bond order potentials) one must convert their Cartesian coordinates into distances and angles. In case of a neural network, a similar step must be performed. The energy contribution of an atom must be invariant with respect to permutations of the same neighbour atoms, to rotations and translations and, possibly, to other symmetry operations. Therefore the Cartesian coordinates are converted to symmetry coordinates which are the actual input to the neural network. The symmetry functions used are somewhat more general than in case of the cluster potentials, but they are also constructed from distances and angles. There are several other approaches as well [15,16]. It is important to realise that the neural network of an atom contains all information of the specific atom type (normally the specific element with nuclear charge Z) and its ‘fingerprint’. This guarantees transferability of the potential between systems of different size and, ideally, chemical composition. Together with the symmetry functions, this also makes up for a conceptionally different viewpoint towards the potential energy of a system, compared with other force fields.

Since the NN potential is analytic, and the atomic contributions sum up, the forces acting on each atom can be calculated without problems by repeated application of the chain rule. The fact that the total energy is the sum of the neural network energy of each atom facilitates also the fitting process, since in quantum chemical calculations the total energy is calculated and individual interaction energies are normally not directly accessible.

Neural network potential energy functions seem to have warranted their existence by being potentially more accurate (but slower) than other force fields and by being faster (but not as accurate) as all-electron (ab-initio) type simulations. Currently, the number of publications in which they are used or developed increases drastically (2016–17: 17, 2014–15: 8, before 2014: 10). Neural network hardware is already built into some processors (for example, Apple’s A11) and dedicated neural network coprocessors like Intel’s Nervana chip are currently entering the market so that the computational aspect looks promising. Altogether it seems to be of interest that such potential energy functions become available in a general-purpose molecular dynamics code like DL_POLY, where they can be combined with the multitude of other interaction types implemented.

2. Implementation details

In the scheme of Behler and Parrinello [14], the total energy is calculated as a sum over individual atomic energies $E_i$. Each contribution $E_i$ depends on the central atom $i$ and its neighbours within a sphere defined by a cutoff distance. The neighbourhood of atom $i$ is described in terms of symmetry functions which map both the relative atomic positions and the atom types into a set of input nodes for the neural network. The number of input nodes is constant and the input node values are independent from exchange of like atoms, rotation and translation. Since the way how energy and forces are constructed varies a bit from other potential energy functions, we have not implemented them as DL_POLY subroutines, but have interfaced the DL_POLY 4.08 code with the neural network library aenet. For a detailed description of the open source atomistic library aenet by Nongnuch and Urban, and the Behler approach in general, we refer to Refs. [14,17–19]. The interface allows to read ANN-parameter files and network structures generated with the aenet-library and to call the aenet energy prediction functions with simple keywords from the DL_POLY input files. It thus adds further functionality to the code, while other functionalities are maintained and can be used in conjunction with the new neural network interface. In particular, it is possible to use the standard analytical functions together with neural network files, which can be useful for repulsive short-range interactions or for long-range interactions that are difficult to cover with the neural network. Also, the virial and the stress tensor are calculated within the new interface, so that the instantaneous pressure is available for simulations performed in a constant pressure ensemble.

When the domain decomposition MPI parallelisation of DL_POLY 4 is used, the total space is divided into spatial
cells which are mapped onto single cores. Each cell stores the information for the atoms within the cell and for the cells neighbourhood within the cutoff including for periodic boundary conditions, the so-called halo. Thus, all relevant information for the call to the neural network library is available for each atom on each node, if the DL_POLY cutoff is chosen larger than any of the cutoffs specified in the neural network. In the current implementation, we forgo a linked-cell algorithm, since extremely large simulations are unlikely to be performed with feedforward neural networks due to their higher computational demand compared to conventional force fields. Instead, a simple neighbour list is updated for every timestep by calculating distances to all neighbours within the domain. Profiling shows that this does not affect the total performance of the interface since the execution time is dominated by the call to the aenet library. A code that supports linked-cell lists is in preparation for the next version of the interface. Having identified all coordinates and types of the neighbours of an atom, its energy contribution \( E_i \) and the forces resulting from the derivative of this energy contribution w.r.t. the positions of atom \( i \) and its neighbours are calculated by the library function \( \text{aenet}_\text{atomic}_\text{energy}_\text{and}_\text{forces} \). This is done for every atom and the energies, forces and virial contributions are summed up, corresponding to the standard implementation for energy prediction with Behler’s method, which we denote with \( \text{nnets} \).

In a second alternative approach, denoted by \( \text{nnpairs} \), the total energy is calculated as a sum over molecular pair potentials. The code identifies all molecule pairs within the cutoff and applies Behler’s method to each pair in turn:

\[
E = \sum_i \sum_{j \neq i} E(i, j),
\]

where \( E_{ij} \) is the pair energy contribution between molecule \( i \) and \( j \) calculated from the sum over atomic energies

\[
E(i, j) = \sum_n E_n(i) + \sum_m E_m(j)
\]

\( n \) goes over all atoms in molecule \( i \) and \( m \) over all atoms in molecule \( j \). For each atomic energy, only the neighbours within the molecular pair are considered and are fed via the symmetry functions into the feedforward neural network. \( \text{nnpairs} \) is slower than \( \text{nnets} \) and loses the ability to describe reactive processes, but allows for more flexibility in terms of combinations of different neural network potentials. Hereby, the fitting is restricted to pairwise interactions between two molecules and each type of molecule-pair is assigned its own neural network. The energies from different molecule-pairwise neural networks are summed up neglecting three-molecule contributions in a similar way that three-atom contributions are neglected in pairwise force fields. For example, it is possible to combine a neural network for water pair interactions with a neural network for sodium-water and one for chloride-water interaction, or for liquid mixtures such as ethylene glycol–water mixtures that are of interest in the framework of solvent dynamics [20]. \( \text{nnpairs} \) may also be an alternative for simulations with larger biomolecules in solution, with neural networks fitting the potential energy surface for each type of molecular pair. It can only be used to describe inter-molecular interactions and flexible molecules need a different treatment for their intra-molecular interactions. Although intra-molecular forces could, in principle, also be treated with a similar neural network, a combination of \( \text{nnpairs} \) with an intra-molecular neural network has not yet been implemented and tested.

3. Liquid water as an example

Given the importance and complexity of systems involving water, it is no surprise that a number of works on molecular dynamics with neural network potentials have already been performed. Small water clusters have been investigated in [21–23]. A number of studies deal with metal or metal oxide surfaces in contact with water [24–26] and with nanosystems and water [27]. If the (auto)dissociation of water in aqueous solutions is to be modelled, reactive potentials are required and such simulations with the neural network approach are reported in [28,29].

In the present work, we use our new neural network implementation to compare two molecular dynamics simulations on liquid water, one with a neural network potential trained from the SPC/E [30] water model, and one using the SPC/E model itself, with otherwise identical conditions. We describe our procedure and compare the results. In other words, we tried to imitate the SPC/E model with the neural network and could directly compare its results with the SPC/E reference. In this example we stay within the pair approximation and do not take advantage of chemical bond breaking/bond formation but strive to verify our implementation which can be readily extended to simulate more complex cases.

At first, a training set has to be created and the network architecture must be specified. The training data was obtained by simulating 1501 collisions of two water molecules with collision energies from 17.3 kcal/mol to 86.48 kcal/mol. Since the SPC/E model consists of pairwise additive terms only, it is not necessary to include data for three or more interacting molecules in the training set. Initial conditions with random relative orientation and impact parameters between 0.0 and 3.0 Å were generated using the venus direct dynamics code [31]. The classical trajectories for non-reactive H\(_2\)O–H\(_2\)O collisions were then obtained with DL_POLY using the standard set of parameters and the water geometry of the SPC/E force field. The collision trajectories were integrated with a time step of 0.5 fs, a total simulation time of 0.75 ps, and a recording stride of 3 for each trajectory. From the whole set, we discarded many of the low-energy configurations with large intermolecular distances and then randomly extracted 10,000 configurations from which 90% were chosen for the training set and 10% for the testing set.

A shallow feedforward neural network with 2 hidden layers and 10 nodes per layer was chosen, together with a hyperbolic tangent activation function. We used a set of Behler type-2 and type-4 symmetry functions [32] for H and for O with a cutoff of 6 Å. The ANN binary parameter files are given in the supplementary information and also contain the chosen parameters of the symmetry functions. The network was trained using the limited-memory Broyden-Fletcher-Goldfarb-Shanno optimisation algorithm [33] up to a final mean average error of
0.4 meV per atom for the testing set. The ANN training was performed with the *aenet* programme [18].

Having thus established a neural network representation of the SPC/E force field, a cubic periodic box of edge length 39.15 Å with 2000 H2O molecules was prepared using packmol [34] and equilibrated with the analytical SPC/E force field for 1 ns. The mass to volume ratio corresponds to a density of 9971 kg/m³. A timestep of 0.2 fs was chosen. The temperature of 298 K was controlled with the Nosé-Hoover thermostat using a relaxation constant of 0.5 ps [35,36]. Two typical trajectories of the x-coordinates of a hydrogen atom are compared in Figure 1 for the full 2 ps simulation time. The trajectories start out at the same point with the same velocity, resemble each other in the first 70 fs and deviate strongly thereafter. This is not unexpected, of course, and it remains to be seen how static and dynamic quantities compare with each other.

First, we discuss static quantities derived from averaging over 2 ps simulations. In Figure 2 and Figure 3, we show the radial distribution functions for oxygen–oxygen distances and oxygen–hydrogen distances respectively. The neural network model can mimic the SPC/E model especially for the short-ranged structure. However, significant differences can be seen for the second coordination shell. On average, a water molecule is involved in $n_{HB} = 3.4$ number of hydrogen-bonds in the SPC/E model, while $n_{HB} = 3.29$ for *nnpairs*. The hydrogen-bond connectivity was obtained in analogy to Ref. [37] with the distance criteria $R_{OO} \leq 3.3$ Å and $R_{OH} \leq 2.35$ Å and the angle criterion $\phi \leq 30^\circ$. Note that we employed a smaller cutoff of 6 Å for the neural network, while we used 10 Å for the original SPC/E model.

Even more interesting is a direct comparison of dynamical quantities. For the rather short 2 ps simulation, we obtain an Einstein diffusion coefficient of $2.95 \times 10^{-9}$ m²/s for SPC/E from the mean square displacements of the water centre of mass, while *nnpairs* reaches a very close value of $2.97 \times 10^{-9}$ m²/s. Both compare well with literature data [38]. As an example of dynamic properties, we calculated the autocorrelation function of the centre of mass velocities, see Figure 4. The *nnpairs* correlation function comes very close to the SPC/E curve and both lead to very
similar power spectra as shown in Figure 5. A discussion on the power spectrum of SPC/E water can be found for example in Ref. [20]. It is interesting to note that nnpairs performs quite good for dynamic quantities although the SPC/E radial distribution in Figures 2 and 3 could not be accurately reproduced for distances larger than the position of the first maximum. A possible reason could be that we used dynamic trajectories of colliding water molecules for the training set generation together with a cutoff of 6 Å which seems to capture the parts of the force field that are needed for dynamics (Figure 5).

In our present case the analytical SPC/E model is faster by a factor of about 100 than the implemented nnpairs model on the same high-performance computing architecture using 4 cores.

4. Conclusions and outlook

An interface between the feedforward neural network library aenet and the general purpose molecular dynamics simulation code DL_POLY has been implemented. The standard implementation nnets follows the Behler approach by considering for each atom all neighbor atoms within a spherical cutoff distance and also allows for reactive simulations. The second approach, nnpairs, uses a pairwise approximation, where energy and forces of molecule-pairs are predicted by the neural network framework. The pair-approximation was tested in comparison to the SPC/E water model reference force field and yielded reasonably accurate results for static and dynamic quantities. Both nnets and nnpairs can be combined with the usual functionality of DL_POLY and in particular also with analytical field and force fields, with rigid body structures, and with thermo- and barostats. It will be interesting to see for example, how well standard analytical intermolecular potentials can be improved by correction of intrinsic inaccuracies with feedforward neural networks, especially in the intermediate distances in between steep repulsive parts of the potential and well known long-range interactions. Our intention now is to make this interface available for the scientific community free of charge.

Acknowledgments

We thank Nongnuch Artrith and Alexander Urban for many fruitful discussions and for sharing their aenet library. The computational results presented have been achieved using the HPC infrastructure LEO of the University of Innsbruck.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This study was supported by the Austrian Science Fund (FWF) project P28979-N27. It was also supported by the Austrian Ministry of Science BMWF as part of the University Infrastructure Program of the Scientific Computing Platform LFU Innsbruck and has received funding from the Euratom Research and Training Programme 2014-2018 under Grant Agreement No. 633053. The authors were also partially supported by the Doctoral Programme on Computational Interdisciplinary Modelling; FP7 Fusion Energy Research.

References


Supplementary Information

A neural network interface for DL_POLY and its application to liquid water

Ivan Sukuba¹², Lei Chen¹, Michael Probst¹, Alexander Kaiser¹*  
¹Institute of Ion Physics and Applied Physics, University of Innsbruck, Technikerstraße 25, 6020 Innsbruck, Austria  
²Department of Nuclear Physics and Biophysics, Faculty of Mathematics, Physics and Informatics, Comenius University, SK-84248 Bratislava, Slovakia

*Corresponding author: alexander.kaiser@uibk.ac.at

The supplementary data is available on


and contains two folders. All files for the training of the neural network with the aenet code [1] and the training set data is stored in “/aenet_training”. The training set data is compressed into “xsf_files_randomsmallset.tar” and contains 10000 files in “xsf” file structure including the configuration energy in eV in the first line. The symmetry functions and its parameters are explicitly given in human-readable form in “01-generate_behler11/” in the two files “H.fingerprint.stp” and “O.fingerprint.stp”. The input files for training and its results are given in “02-train_behler11”. The final neural network parameters including also the symmetry function information are stored in the binary files “H.10t-10t.ann” and “O.10t-10t.ann” and are directly used by the DL_POLY interface in the folder dl_poly_run.

Iterative training set refinement enables reactive molecular dynamics via machine learned forces†

Lei Chen, a Ivan Sukuba, ab Michael Probstac and Alexander Kaiser a,*

Machine learning approaches have been successfully employed in many fields of computational chemistry and physics. However, atomistic simulations driven by machine-learned forces are still very challenging. Here we show that reactive self-sputtering from a beryllium surface can be simulated using neural network trained forces with an accuracy that rivals or exceeds other approaches. The key in machine learning from density functional theory calculations is a well-balanced and complete training set of energies and forces. We have implemented a refinement protocol that corrects the low extrapolation capabilities of neural networks by iteratively checking and improving the molecular dynamic simulations. The sputtering yield obtained for incident energies below 100 eV agrees perfectly with results from ab initio molecular dynamics simulations and compares well with earlier calculations based on pair potentials and bond-order potentials. This approach enables simulation times, sizes and statistics similar to what is accessible by conventional force fields and reaching beyond what is possible with direct ab initio molecular dynamics. We observed that a potential fitted to one surface, Be(0001), has to be augmented with training data for another surface, Be(0110), in order to be used for both.

Introduction

In molecular dynamics (MD) simulations energies and forces are complicated functions of nuclear coordinates and element types. Calculating forces on-the-fly by electronic structure methods avoids having to handle these functions explicitly but is, even with density functional methods, still restricted to small systems and short times, compared to MD with analytic potential energy functions. Even for non-reactive systems the development of a reliable force field is very tedious. Consequently, machine learning approaches are developed to fill this gap by learning energy and forces from quantum chemical data and to replace a conventional force field.1–3 Feedforward Neural Networks4,6–8 and Gaussian Approximation Potentials9–12 are most widely used at present. In both of them Cartesian coordinates of the atoms are first transformed into symmetry invariant atom-centered representations by various methods.7,13 The present work applies feedforward neural networks and the Behler-Parrinello type atomistic representation. The parameters are the bias parameters that act on individual neurons and the weights that interconnect artificial neurons in different layers. The number of parameters only depends on the size of the neural network. When increasing the number of parameters, the network can basically store the information it is trained on almost perfectly and it can also interpolate to some extent.

We can train the network on a finite set of energies and forces and its quality will depend very much on the choice and size of the training set and on the power of the global optimizer to reach a low-lying minimum. Compared to the huge positional phase space spanned by all combinations of atomic positions, the number of configurations (~6000 in this work) used to train the network is meagre.

In this work, we implement a refining procedure based on previous work4,14 for training set construction and we train a neural network potential (NNP) for molecular dynamics simulations of reactive beryllium (Be) self-sputtering and show that our NNP based simulations are accurate in predicting sputtering yields. Knowledge of the stability of Be surfaces is very relevant because beryllium sheets have been chosen as armor material in the first wall of the ITER reactor currently being constructed.15 Having only 4 valence electrons, beryllium has the additional advantage that it can be treated with density functional calculations rather efficiently, making comparisons with ab initio MD feasible. Previous work on this system include an MD study by Ueda et al. where a pair potential was developed and self-sputtering processes of Be at low incident energies (≤100 eV) was simulated.16 Björkas et al. developed a bond order potential for the ternary system Be–C–H, and the Be potential was applied to MD simulations of Be self-sputtering.17
Training set generation and refinement

The NNP is trained on energies and forces obtained with plane-wave DFT calculations. Details of the DFT and ab initio MD calculations are given in the section on Computational methods. Atrith and Behler have already described a refinement procedure based on dynamic simulations to extend the accuracy and applicability of a neural network potential for MD simulations.4,14 In the present work, our refining procedure also relies on the assumption that two different neural networks that have been fitted to the same data set will deliver approximately the same result for well-sampled regions of the phase space but not in extrapolations out of these regions. This allows to systematically and automatically identify structures that are missing in the training data. The iterative procedure is schematically shown in Fig. 1. In a first step, configurations were created by randomly extracting snapshots from 500 ab initio MD sputtering trajectories on a small Be(0001) surface slab with 96 atoms. Two preliminary NNP’s, NNP1 and NNP2 were fitted to this training set. NNP1 and NNP2 have the same topology and differ only in the starting values of their fit parameters which are randomly chosen. They are simple feedforward $3 \times 30 \times 30 \times 1$ NNs with two hidden layers. We

found that including more parameters or making the neural network deeper does not improve the accuracy any more. More details on the employed symmetry functions and the neural network are given in the Computational methods.

The refinement procedure starts with using NNP1 for short MD simulations (40 fs) of the Be self-sputtering process at various impact energies on the surface slab with 96 atoms. The energies and forces of these new configurations were then predicted by NNP2 along the same trajectories, and energies and forces of both networks were compared with each other as shown in Fig. 2. Configurations with energy differences larger than $\epsilon_E = 20$ meV per atom or maximum force differences larger than $\epsilon_F = 2$ eV Å\(^{-1}\) were selected and subjected to a DFT calculation of energies and forces which were then added to the training data.

Two new neural network potentials NNP3 and NNP4 were fitted to the refined training set and their differences for new trajectories is shown in Fig. 3. It is apparent that already after one cycle of the refinement process, the differences between the two NNPs decrease considerably. For most configurations in Fig. 3, the energy and maximum forces differences between NNP3 and NNP4 are within 5 meV per atom and 1 eV Å\(^{-1}\). Due to this excellent improvement, we reduced the number of trajectory calculations in the second refinement step.

Further refinement can be done iteratively as indicated in Fig. 1. In our case, a second refinement step with much smaller energy and force thresholds ($\epsilon_E = 2$ meV per atom, $\epsilon_F = 0.8$ eV Å\(^{-1}\)) was sufficient.

With this iterative refinement process, the final reference data set consists of 5871 configurations containing 97 atoms

---

**Fig. 1** Schematic of the iterative refinement procedure based on previous work4,14 for training set generation. Energies and forces predicted with NNP1 and NNP2 are compared with each other along molecular dynamics trajectories generated with NNP1. The decision (NNP1$\Rightarrow$NNP2?) of including a particular phase-space point into the training set is made either by the energy criterion $|E_{\text{NNP1}} - E_{\text{NNP2}}| > \epsilon_E$ or the force criterion $|F_{\text{NNP1}} - F_{\text{NNP2}}| > \epsilon_F$.

**Fig. 2** Comparison of energies and forces of the preliminary NN potentials NNP1 and NNP2 along the same trajectories. Configurations with absolute values of energy differences higher than $\epsilon_E = 20$ meV per atom or maximum force differences higher than $\epsilon_F = 2$ eV Å\(^{-1}\) have been recalculated by DFT and added to the training data to obtain the refined NNP in Fig. 3.
Each. Thus, 5270 energies and 1,533,486 forces are used to fit the final potential energy function NNP5 and 601 energies and 174,870 forces are part of the test set which is not used for training but rather to validate the potential and to prevent overfitting. After 60 training epochs, the root mean square errors (RMSE) in the test set converged to 0.7 meV per atom for energies and 32.0 meV Å⁻¹ per atom for forces, very close to the corresponding value in the training set (0.6 meV per atom and 31.8 meV Å⁻¹).

Static performance of the refined neural network potential

The correlation of NNP5 and DFT energies and forces for each atom are shown in Fig. 4(a) and (b). Only the x-component of the forces is shown here since y and z have been inspected but give virtually identical plots. DFT and NNP5 energies are very close except for very few configurations in the training set. Similarly, the values of NNP5 and DFT forces at all three directions are perfectly correlated.

We are now in a position to go to a larger system, a Be(0001) surface with 490 atoms. This surface will later be used for the sputtering simulations. The equilibrium lattice constants of using NNP5 turn out reasonable (Table 1) with NNP5 and DFT showing relative differences of less than 2% compared to experimental values. The NNP5 total energy for this configuration is also very close to the DFT value with an absolute deviation of 3.3 meV per atom. Relaxation of the surface to its equilibrium configuration is necessary for the subsequent sputtering calculations since otherwise, a large amount of potential energy is heating the system at the beginning that could irreversibly change the structure due to expansion and melting.

Performance of the refined neural network potential in reactive sputtering simulations

Although only a small system was used to train NNP5, the energy contribution from each atom depends only on the local chemical environment and therefore it can be used to simulate a larger system. MD simulations performed on the Be(0001) surface slab with 490 atoms result in a self-sputtering yield of 5.6% which agrees perfectly with a yield of 5% obtained from 500 ab initio MD trajectories calculated on the same system under identical conditions for 100 eV incident energy. The sputtering yields of all our simulations
are summarized in Table 2 and compared with other MD simulations, Monte Carlo data and experimental estimates in Fig. 5. Statistically, their accuracy increases with the number of simulation runs and the error bars are estimated by $D = 2\sigma/\sqrt{N}$ where the standard deviation $\sigma$ is obtained by assuming a Bernoulli distribution of $N$ trajectories. The NNP5 sputtering yields fall well within other available data. Timings included in Table 2 show that the NNP5-based MD simulations are more than two orders of magnitude faster than their \textit{ab initio} MD simulations.

For all incident energies, we never encountered reflections of the Be projectile. In our energy range this is expected and agrees with the simulation results from Ueda \textit{et al.}\cite{21} where a pair potential had been used. No atoms were sputtered for incident energies below 50 eV, which is consistent with the findings from Ueda \textit{et al.} only if the large error bars are taken into account but not in good agreement with the BOP-based results of Björkas \textit{et al.} who reported a similar value than Ueda \textit{et al.} at 50 eV incident energy but with a much smaller statistical error.\cite{16} The sputtering threshold energy defined by no sputtering event occurring in 500 trajectories lies within 50 and 55 eV for our setup with NNP5 and is higher than previous estimates from 16–25 eV.\cite{17,19}

We used an estimated sputtering threshold energy $E_{\text{th}} = 53$ eV, the parameters $q = 0.82$, $\mu = 1.34$ and $\lambda = 2.03$ from literature\cite{19} and the experimental data at high energies to fit the sputtering yields to the Eckstein formula.\cite{20} The resulting function is also included in Fig. 5. Since the available experimental data is in the keV range, it is not possible to make a direct comparison with our simulated results. At an incident energy of 100 eV, our result is very close to the values from Roth \textit{et al.},\cite{22} Ueda \textit{et al.}\cite{21} and Björkas \textit{et al.}\cite{16} At 75 eV, the NNP5 based sputtering yield is very close to that from the bond-order potential.\cite{17}

In order to check the convergence of our model system with respect to surface size, we also simulated a system with 2000 atoms (the crystal size is given in Table 2) at an incident energy of 100 eV. With 500 simulation runs, we obtained 27 sputtering events (sputtering yield 0.054) and obtained very good agreement with the smaller system (sputtering yield 0.056).

### Transferability of the neural network potential

With the purpose of testing the transferability of NNP5 to a different surface structure that had not been included in the training set, we performed self-sputtering simulations with an incident energy of 75 and 100 eV on a Be(0110) surface consisting of 480 atoms. We obtained much smaller sputtering yields than reported by Ueda.\cite{21} Applying an iterative refinement step as described above on Be(0110) and refining the neural network, more reasonable results are obtained, albeit of course now with a different potential (NNP6). The sputtering yields for the Be(0110) surface using NNP6 are summarized in Table 3 and plotted in Fig. 6. A fit to the Eckstein formula is also shown. We used the same values of the parameters $q$, $\mu$ and $\lambda$ as for the Be(0001) surface but a lower estimate of the threshold energy $E_{\text{th}} = 30$ eV.\cite{22} The sputtering yield at 100 eV for NNP6 is much larger than the one from NNP5 and is comparable to Ueda’s results.\cite{21} The sputtering yields for the (0001) and (0110) surfaces are close to each other at 100 eV, while at lower energies the (0110) surface is more susceptible to sputtering with the simple reason that the Be(0001) surface is more stable. In fact, the DFT calculated surface binding energy of 5.13 eV for the (0001) surfaces is much higher than 2.48 eV for the (0110) surface.\cite{22} Finally, we note that the upgrade from NNP5 to NNP6 conserves the accuracy for the (0001) surface with a sputtering yield of 0.052 (0.056) for NNP6 (NNP5) on 500 trajectories.

<table>
<thead>
<tr>
<th>Crystal size (Å)</th>
<th>Sputtering yield of Be(0001)</th>
<th>50 eV</th>
<th>55 eV</th>
<th>60 eV</th>
<th>65 eV</th>
<th>75 eV</th>
<th>100 eV</th>
<th>CPU time/trajectory</th>
</tr>
</thead>
<tbody>
<tr>
<td>$15.6 \times 15.6 \times 30.1$ (490 atoms)</td>
<td>NNP5</td>
<td>0</td>
<td>0.003 (0.0015)</td>
<td>0.0072 (0.0024)</td>
<td>0.0086 (0.0026)</td>
<td>0.026 (0.014)</td>
<td>0.056 (0.021)</td>
<td>10 minutes (4 cores)</td>
</tr>
<tr>
<td>ab initio MD</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.050 (0.019)</td>
<td>30 hours (16 cores)</td>
<td></td>
</tr>
<tr>
<td>$22.8 \times 22.8 \times 48.4$ (2000 atoms)</td>
<td>NNP5</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>0.034 (0.020)</td>
</tr>
</tbody>
</table>

### Table 2: Sputtering yields and statistical error estimates for the Be(0001) surface obtained from neural network based MD simulation at various incident energies and from an \textit{ab initio} MD simulation at 100 eV incident energy. The average computational time for each NNP based MD and \textit{ab initio} MD simulation is compared. Note that we have performed 5000 trajectories for low incident energies (55 eV, 60 eV and 65 eV) using NNP5 to obtain a lower error bar.
Summary
The self-sputtering simulations based on the refined neural network potential give promising results for small and large periodic cell-sizes of the Be(0001) surface. The sputtering yields agree with full \textit{ab initio} Born–Oppenheimer MD results at 100 eV incident energy. Compared to literature data, our simulation protocol seems to be quite competitive while not requiring manual potential development. Transferability to another surface direction was only possible by additional refinement and including data of the new surface into the training data. Although the neural network can only be used for scenarios that are included in the training data and this involves thousands of \textit{ab initio} single-point calculations, the final neural network potential can be used on longer time-scales and larger systems. Especially, there is no difference conceptually in treating more complicated systems such as alloys, where the construction of conventional force fields becomes increasingly cumbersome. As a next step, we plan to apply such simulations to other plasma–wall interactions, in particular considering the ternary H–Be–W system.

Computational methods

Neural network potential
In this work, we train a neural network potential of the Behler–Parrinello type.\textsuperscript{6} The neural network code \textit{n}2\textit{p}2 (ref. 25 and 26) recently developed by A. Singraber \textit{et al.} is based on Behler’s work.\textsuperscript{6,7} It includes both force and energy fitting and implements a Kalman filter optimizer\textsuperscript{26} which can deal with the large number of data points when forces are included in the fitting. It has also been linked to the versatile molecular dynamics code LAMMPS that we used to integrate the sputtering trajectories.\textsuperscript{27,28} In the Behler–Parrinello approach,\textsuperscript{4} the total energy $E_{\text{tot}}$ of one configuration is the sum of atomic energies $E_i$ provided by element-specific neural networks that depend on the local atomic neighbourhood only. Atomic coordinates are transformed to symmetry-invariant atom-centered symmetry functions before entering the neural network in the input nodes. Efficient fitting to forces requires analytic gradients implemented in the \textit{n}2\textit{p}2 library. We chose a simple feedforward neural network topology with two hidden layers with 30 neurons each. The so-called soft-plus activation function\textsuperscript{29,30} which is a smooth approximation of rectified

Table 3  Calculated sputtering yields and statistical error estimates for the Be(0110) surface obtained from neural network based MD simulations at various incident energies

<table>
<thead>
<tr>
<th>Crystal size (Å)</th>
<th>Sputtering yield of Be(0110)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>20 eV</td>
</tr>
<tr>
<td>13.7 \times 17.7 \times 28.8</td>
<td>NNP5</td>
</tr>
<tr>
<td></td>
<td>NNP6</td>
</tr>
<tr>
<td></td>
<td>Ueda \textit{et al.}\textsuperscript{16}</td>
</tr>
</tbody>
</table>
linear units (RELU) was used as recommended. A cutoff radius of 7 Å is sufficient to include all relevant neighbour atoms. The input consists of 9 radial, 24 angular narrow and 20 angular wide Behler-type symmetry functions as detailed in the ESI.†

Density functional theory

The static DFT and the ab initio MD simulations to generate training data were carried out using the Vienna Ab initio Simulation Package (VASP). The core and valence electrons were described by the Projector Augmented Wave (PAW) method and the Perdew–Burke–Ernzerhof (PBE) exchange–correlation functional. A plane wave basis set with a cut-off energy of 350 eV with periodic boundary conditions was used. The PAW potential for beryllium was used as provided in the VASP library. A Gamma-centered k-point mesh of 3 × 3 × 3 was employed. The initial training set was generated by performing Born–Oppenheimer ab initio MD on a small hexagonal closely packed Be surface (0001) with 96 atoms (9.1 × 9.1 × 20.6 Å). It was first relaxed with a convergence criterion of 10−5 eV on the total energy (about 10−4 meV per atom) and of 1 meV Å−1 on the forces. Subsequently, the relaxed surface was simulated at 300 K for 2 ps within the canonical ensemble using the Nosé–Hoover algorithm. Then perpendicular impacts of single Be atoms with energies of 20, 35, 50, 75 and 100 eV starting from a distance of 5 Å above the surface were simulated. 100 ab initio MD runs were performed for each impact energy. The time step was chosen to be 0.5 fs and one run lasted 150 fs for low impact energies (20, 35, 50 eV). Impacts with energies of 75 and 100 eV were simulated for 50 fs.

MD simulations of sputtering on neural network potentials

In our MD simulations of non-accumulative self-sputtering, an incident neutral Be atom impacts on a pristine Be surface. The target consists of 490 atoms with a size of 15.6 × 15.6 × 30.1 Å. Its crystal structure was relaxed and equilibrated for 2 ps at 300 K within the NVT ensemble using the Nosé–Hoover thermostat before running the trajectories. The incident particle was initially placed 5 Å above the surface while its x and y coordinates were randomly chosen. Kinetic energies of 50, 55, 60, 65, 75, and 100 eV were assigned to it by a respective initial velocity in z-direction, thus only impacts perpendicular to the surface were simulated in this work. Trajectories were initially integrated for 120 fs with an integration step of 0.2 fs. Further 120 fs integration time were added to those trajectories where the decision of an observed sputtering event could not be made after the first 120 fs. 500 separate MD runs were performed for each incident energy. For the MD simulations in the training data refinement process, the computational details are identical to the sputtering simulations for larger surface systems except for a shorter integration time of 40 fs.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

The authors are grateful to Andreas Singraber for his valuable suggestions concerning the NNP fitting. This work was supported by the Austrian Science Fund (FWF): P28979, Tyrolean Science Fund (TWF): UNI-0404/2308 and the Kommission zur Koordination der Kernfusionsforschung der ÖAW (KKKÖ) and has partially been carried out within the framework of the EUROfusion Consortium. We have received funding from the Euratom research and training programme 2014–2018 and 2019–2020 under grant agreement no. 633053. The views and opinions expressed herein do not necessarily reflect those of the European Commission. The computational results presented have been calculated using the Vienna Scientific Cluster (VSC-3) and the HPC-infrastructure LEO of the University of Innsbruck.

Notes and references

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt durch meine eigenhändige Unterschrift, dass ich die vorliegende Arbeit selbständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe. Alle Stellen, die wörtlich oder inhaltlich den angegebenen Quellen entnommen wurden, sind als solche kenntlich gemacht.

Die vorliegende Arbeit wurde bisher in gleicher oder ähnlicher Form noch nicht als Magister-/Master-/Diplomarbeit/Dissertation eingereicht.

Datum

Unterschrift