Title
Self-limited Etching for Atomic Scale Surface Engineering of Metals: Understanding and Design

Author
Xia, Yantao

Publication Date
2023

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA
Los Angeles

Self-limited Etching for Atomic Scale Surface Engineering of Metals: Understanding and Design

A dissertation submitted in partial satisfaction
of the requirements for the degree
Doctor of Philosophy in Chemical Engineering

by

Yantao Xia

2023

ABSTRACT OF THE DISSERTATION

Self-limited Etching for Atomic Scale Surface Engineering of Metals: Understanding and Design

by

Yantao Xia
Doctor of Philosophy in Chemical Engineering
University of California, Los Angeles, 2023
Professor Philippe Sautet, Chair

Plasma-thermal atomic layer etching (ALE) is an emerging direct metal etch (DME) method that can potentially enable integration of many metals previously considered impossible to pattern. By combining a low-temperature plasma activation with thermal removal, metal layers can be patterned directionally and selectively. However, little is understood about the surface chemistry that leads to directionality and selectivity. In this dissertation, a suite of modeling tools is developed, ranging from thermodynamic screening of potential process chemistries, to development of accurate interatomic potential models, large-scale atomistic simulation of plasma processes, and finally the analysis of trajectories. The thermodynamic screening model is applied to combinations of oxygen/nitrogen plasma activation on nickel/-copper surfaces, using formic acid/formamidinate as the etchant. In total 8 processes were screened computationally, predicting a nitrogen-plasma based process on nickel metal etch will yield similar etching characteristics as previously demonstrated oxygen-plasma based process. On all combinations of modifier/substrates it is predicted that higher surface coverage generally leads to more favorable etching, and there exists a critical coverage below which etching is unfavorable thermodynamically. It is found that inserting ions into the subsurface sites (possible through the impact ion energy) makes etching highly favorable. This protocol can be readily extended to other combinations of metals/modifiers/etchants to allow for a

rapid screening of etching chemistries. The demonstrated complex site-dependence of the etching energetics is accounted for explicitly in molecular dynamics simulations enabled by a machine learning interatomic potential for copper and oxygen trained on *ab-initio* calculation data. A large scale simulation protocol for atomisitc plasma-surface interaction simulation is developed, and used to obtain atomically-resolved trajectory of copper oxidation. The simulation results show that for a low-energy plasma (kinetic energy $\leq 20\,\mathrm{eV}$), the ions do not penetrate into the substrate lattice. The oxidation process in the bulk of the film is still diffusion limited. The effect of ion energy lies in delivering additional, depth-dependent thermal energy that promotes diffusion within the oxide film. It is confirmed that at $80\,^{\circ}\mathrm{C}$ the oxidation is not self-limiting. The chemical identity of the oxide film is determined to be mainly CuO. The crystalinity is further studied in a separate set of simulations that increased ion flux 4-fold, effective accelerating simulation. Repeated, prolonged ion-impact on the already-oxidized film (2 nm thick oxide) leads a layer of crystalline CuO beneath an atomically-thin, but rough top layer of amorphous CuO. This is not observed in other non-accelerated simulations, which always gave amorphous CuO structures. Presumably this is due to crystallization being a slow process dominated by rare events. By investigating the effect of process-relevant parameters (temperature, ion-to-neutral ratio, and ion kinetic energy), it is found that self-limited growth may be possible by lowering the subtrate temperature. In such conditions, the limiting thickness is controllable by tuning the ion energy distribution function (IEDF) in the plasma. This is due to the limited range of energy delivery through collision cascade in the oxide film.

The dissertation of Yantao Xia is approved.

Jane Pei-Chen Chang

Panagiotis D. Christofides

Daniel Neuhauser

Philippe Sautet, Committee Chair

University of California, Los Angeles

2023

To workers of the world

Contents

List of Figures

List of Tables

Preface

The author gratefully acknowledges the support from UCLA Institute of Digital Research and Education(IDRE) for computational resources on the UCLA Hoffman2 cluster.

This work was funded by NSF, CBET division, Grant No. 2212981.

This work used the the Extreme Science and Engineering Discovery Environment (XSEDE) Comet and Expanse clusters at San Diego Supercomputer Center, Bridges2 cluster at Pittsburgh Supercomputing Center through allocation TG-CHE170060. XSEDE is supported by National Science Foundation Grant number ACI-1548562.

Results from "Xia Sang, Yantao Xia, Philippe Sautet, and Jane P. Chang "Atomic layer etching of metals with anisotropy, specificity, and selectivity", Journal of Vacuum Science & Technology A 38, 043005 (2020);" is reproduced Chapter 4. This work is produced in collaboration with Xia Sang (for the experimental part) and his contribution is acknowledged here.

Results from "Yantao Xia and Philippe Sautet "Thermodynamics of Atomic Layer Etching Chemistry on Copper and Nickel Surfaces from First Principles", Chemistry of Materials, 33, 17, 6774-6786 (2021);" is reproduced Chapter 3.

Results from "Yantao Xia and Philippe Sautet "Evolution of Copper Surfaces under Plasma Oxidation:Molecular Dynamics Study with Neural Network Potentials", ACS Nano, 16, 12, 20680–20692 (2022)" is reproduced Chapter 5 and Chapter 6.

Curriculum Vitae

2014 – 2018	B.S. in Chemical Engineering, University of California, Los Angeles (UCLA), Los Angeles, California.
2018 – 2020	M.S. in Chemical Engineering, University of California, Los Angeles (UCLA), Los Angeles, California.
2020 – Present	Ph.D. Student in Chemical Engineering, University of California, Los Angeles (UCLA), Los Angeles, California.

Publications

Huan Wang, Sijia Guo, Haiou Zeng, Chenhao Zhang, **Yantao Xia**, Chenghao Weng, Yichen Deng; Luda Wang, "Glycolysis maintains AMPK activation in sorafenib-induced Warburg effect", Bioengineering & Translational Medicine, 2023, *Submitted*

Yantao Xia and Philippe Sautet "Evolution of Copper Surfaces under Plasma Oxidation: Molecular Dynamics Study with Neural Network Potentials", ACS Nano, **16**, 12, 20680–20692 (2022)

Yantao Xia and Philippe Sautet "Thermodynamics of Atomic Layer Etching Chemistry on Copper and Nickel Surfaces from First Principles", Chemistry of Materials, **33**, 17, 6774-6786 (2021);

Yantao Xia M.S. Thesis, Thermodynamic Screening of Reaction Chemistries for Atomic Layer Etching of Metals, UCLA, 2020.

Xia Sang, **Yantao Xia**, Jane P. Chang, and Philippe Sautet "Directional and selective patterning of EUV absorbers", Proc. SPIE 11615, Advanced Etch Technology and Process

Integration for Nanopatterning X, 1161503 (2021);

Xia Sang, **Yantao Xia**, Philippe Sautet, and Jane P. Chang "Atomic layer etching of metals with anisotropy, specificity, and selectivity", Journal of Vacuum Science & Technology A **38**, 043005 (2020);

Samuel Hernandez, **Yantao Xia**, Vojtěch Vlček, Robert Boutelle, Roi Baer, Eran Rabani, and Daniel Neuhauser "First-principles spectra of Au nanoparticles: from quantum to classical absorption", Molecular Physics, 116:19-20, 2506-2511

CHAPTER 1

Research Background

All solids terminate by a surface. As a result, surface processes are key for a vast array of applications. For the ultimate control of the composition and structure of surfaces, selective and accurate engineering methods of surfaces at the nanometer/subnanometer scale are required. This is especially the case for the microelectronics industry, which, following the trend predicted by Moore's Law, has entered the nanoelectronics era. As of 2023, critical dimensions (i.e. "feature length") in major producers have entered the single-digit nanometer scale.[125] Specifically, the precise control of film thickness is becoming necessary with applications such as extreme ultraviolet (EUV) lithography masks and 3D devices. This is in addition to the already stringent requirement of patterning planar features.[104]

An overview of the photolithographic pattern transfer process is depicted in figure 1.1. A photosensitive polymeric material, called the photoresist (PR), is first coated onto the substrate to be patterned. The resist is then exposed via a photomask. In the case depicted here (for positive PR), the exposed region becomes soluble in the *developer* solution and is removed from the substrate during development. Now the pattern is transferred from the photomask to the photoresist. In the etching step, the photoresist acts as a mask for etching. Only the exposed parts of the substrate would be etched, transferring the pattern onto the substrate."

The focus of the present dissertation is etching. Conventional wet etching approaches used in the fields of metallurgy and art expose patterned metals to various acids. The metal becomes oxidized and dissolves into the liquid phase. Wet etching offers fast etch speed but the Brownian motion of the etchant molecules always results in uniform etching in all directions (isotropic etching), rendering it unsuitable to pattern the high aspect ratios

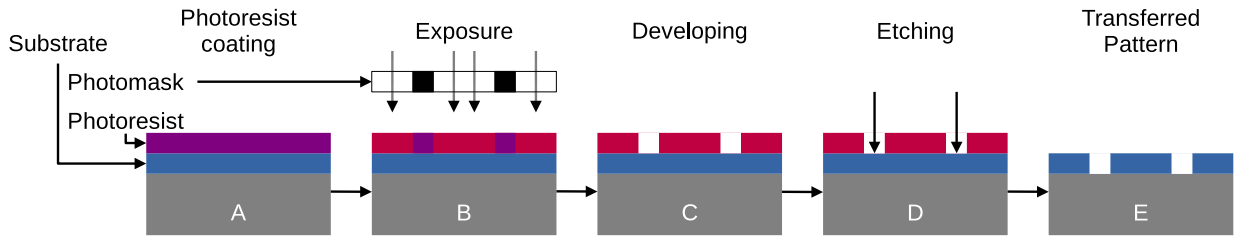


Figure 1.1: **A**: photoresist (PR) is coated onto the substrate. **B**: The exposure via a photomask transfers the pattern from the mask to the PR. **C**: The exposed regions dissolve in the developer solution and expose the target region on the substrate. **D**: These regions are then brought into contact with an etching environment. **E**: The resist is then ashed, leaving the transferred pattern on the substrate.

structures today. In addition, the harsh chemical conditions may be corrosive to existing structures.

Today, dry etching is the dominant technique for patterning. Two mechanisms are at work in dry etching: (1) chemical etch process, in which neutral radicals react with the surface, producing volatile products, and (2) sputtering, which relies on the kinetic energy of inert ions to disrupt the lattice and eject substrate atoms. Common dry etching techniques are shown in figure 1.2. Neutral chemical etch (**A**) is similar to wet etching, is always isotropic, and able to reach underneath the masking material. In sputtering (**B**), the ions are accelerated toward the substrate, leading to vertical direction being preferred, but the side wall slopes outward due to the finite mean free paths of the ions, making it difficult to create high-aspect-ratio structures. In reactive ion etching (RIE, **C**), nearly vertical side walls are achieved when chemical etch is re-introduced. In this technique, mechanisms (1) and (2) cooperate. Structures with still higher aspect ratios are possible with inhibitor films in ion-enhanced inhibitor etching (**D**), where the sidewalls are protected from etchant contamination by repeated deposition of the polymeric inhibitor film.

A major drawback of the aforementioned existing continuous dry etching techniques is their inability to precisely control vertical etch thickness, which provides the opportunity window for atomic layer etching (ALE).

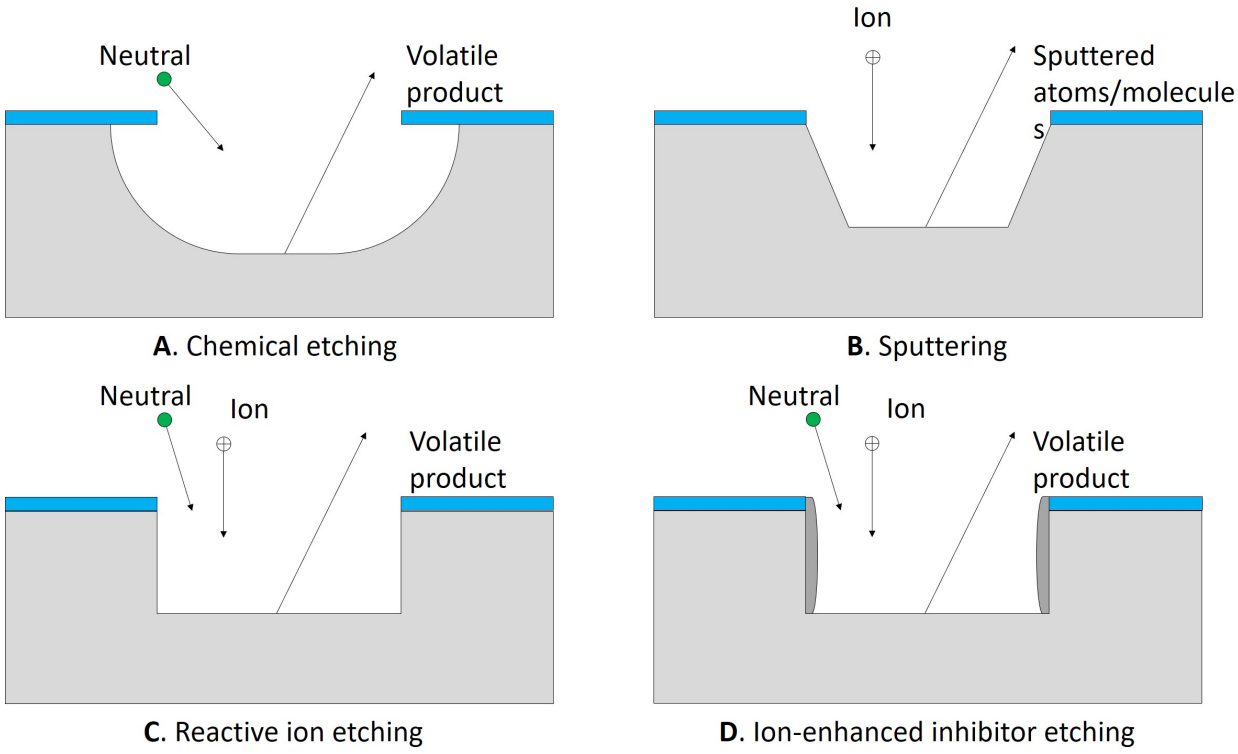


Figure 1.2: Schematic of different vertical etching profiles. **A)** neutral radicals in chemical etching produce isotropic profiles. The substrate underneath the mask is etched away. **B)** sputtering is anisotropic, but cannot create deep trenches due to the non-vertical sidewall. **C)** Ion-enhanced etching is able to create near-vertical sidewalls. **D)** Inhibitor etching protects the sidewall by an inhibitor, enabling even higher aspect ratio and prevents excessive sidewall damage from neutrals (reactive radicals).

The conceptual reverse of atomic layer deposition (ALD), ALE was first reported in the 1980s. While ALD has since seen wide-range industry adoption, ALE was considered too slow to be practical until developments in the last decade brought it back under intense investigation when applications emerged where precision rather than speed is the bottleneck.[108] The technique can be thought of as RIE with time-separated doses. In RIE, both ions and radicals are present at all times. In ALE, the ions and radicals are introduced in time-separated pulses. The ions are used to modify the surface, creating a surface layer of different chemical composition, but do not themselves sputter etch. The neutrals are introduced after ions are purged. The removal is accomplished either by sputtering with inert species (e.g., Ar) or a reaction that forms a volatile compound. Under both scenarios, only the activated pattern is removed, and material removal stops when the pristine substrate is exposed. In both steps, the reaction limits itself to the surface layers. In contrast to RIE, in ALE the energy of the ions must be kept as low as possible to avoid sputtering and/or diffusion deep inside the substrate. The self-limiting nature of the reaction enables digital thickness control, as the etch rate no longer depends on exposure time.

The main engineering challenge in ALE is that surface reactions must be mutually exclusive. Consider the plasma ALE process on Si,[110] depicted in figure 1.3. In the activation step, plasma-assisted chlorine adsorption activates the Si surface. The energy source for the removal step is Ar^+ plasma. In the activation step, the ion energies must be kept low to stay in the ALE regime. In this case, the process window that allows for ALE is 40 V–60 V in terms of RF bias in the chlorination step. Higher energies result in a loss of selectivity due to spontaneous etch. Etching at lower energies is too slow. At the middle of this range, the etch-per-cycle (EPC) is approximately 12 Å, corresponding to roughly 3 layers of atoms.

The fact that the same chemistry can be used in both ALE and continuous modes enables an interesting comparison. If Cl_2 and Ar doses are not time-separated, a continuous etch results. The advantage of a low-energy process is seen through the Si to SiO_2 selectivity: selectivity of ALE is an order of magnitude higher than that of RIE within the ALE operating window.[110]

The requirement of self-limiting surface reactions can be examined in more detail given

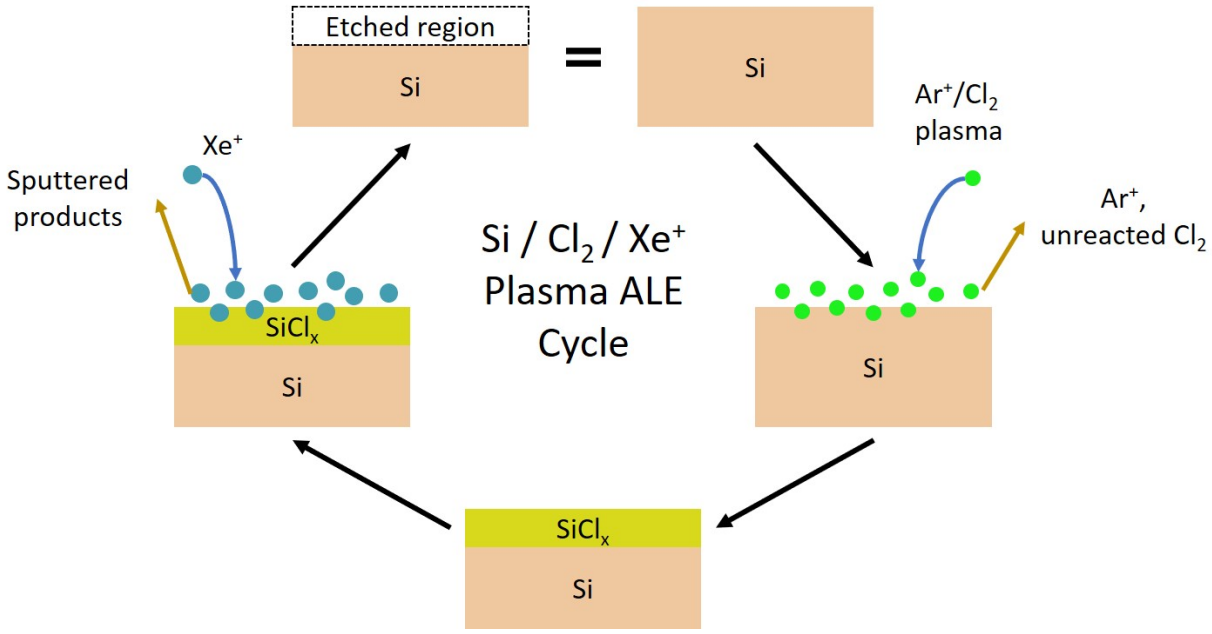


Figure 1.3: $Si/Cl_2/Ar$ plasma ALE Process. Chlorination-activated Si surface is sputter-etched with Ar^+ plasma.

this context. Here, the chlorine activation must be self-limiting because if Cl atoms diffuse spontaneously into the Si lattice, the thickness of the $SiCl_4$-like layer formed (and subsequently etched) would be dependent on Ar^+ exposure time. Similarly, Ar^+ plasma must not be energetic enough to sputter the Si substrate on its own.

The $Si/Cl_2/Ar^+$ example is entirely plasma-based. The alternative, thermal ALE, is discussed in the next example. Consider the process depicted in figure 1.4. HF gas adsorbs on the surface of alumina (Al_2O_3) to form a stable and nonvolatile AlF_3-like layer, releasing H_2O as a by-product. After purging HF, the AlF_3 layer is reacted with trimethylaluminum (TMA) to form the volatile $AlF(CH_3)_2$ complex that must be purged before the HF is re-introduced. The dangling methyl groups on the surface are removed at the next cycle by HF to form methane. In addition to the same requirement that fluorination does not proceed indefinitely and that TMA is not reactive on Al_2O_3, the case here is complicated by the fact that TMA and HF are used for ALD of AlF_3 at lower temperatures. The ALE reaction can

be written overall (unbalanced) as follows:

$$AlO_x(CH_3)_{3-x} + HF \longrightarrow AlF_3 + H_2O\uparrow + CH_4\uparrow \quad \{1.1a\}$$

$$AlF_3 + Al_2O_3 + Al(CH_3)_3 \longrightarrow AlF(CH_3)_2\uparrow + AlO_x(CH_3)_{3-x} \quad \{1.1b\}$$

The competing ALD reaction is as follows:

$$AlF_x(CH_3)_{3-x} + HF \longrightarrow AlF_3{-}HF + CH_4\uparrow \quad \{1.2a\}$$

$$AlF_3{-}HF + Al(CH_3)_3 \longrightarrow AlF_x(CH_3)_{3-x}\downarrow + CH_4\uparrow \quad \{1.2b\}$$

The competition between ALE of Al_2O_3 and ALD of AlF_3 makes the time separation of doses critical, as the simultaneous presence of the etchant HF and the etching product $Al(CH_3)_3$ leads to the deposition of AlF_3.

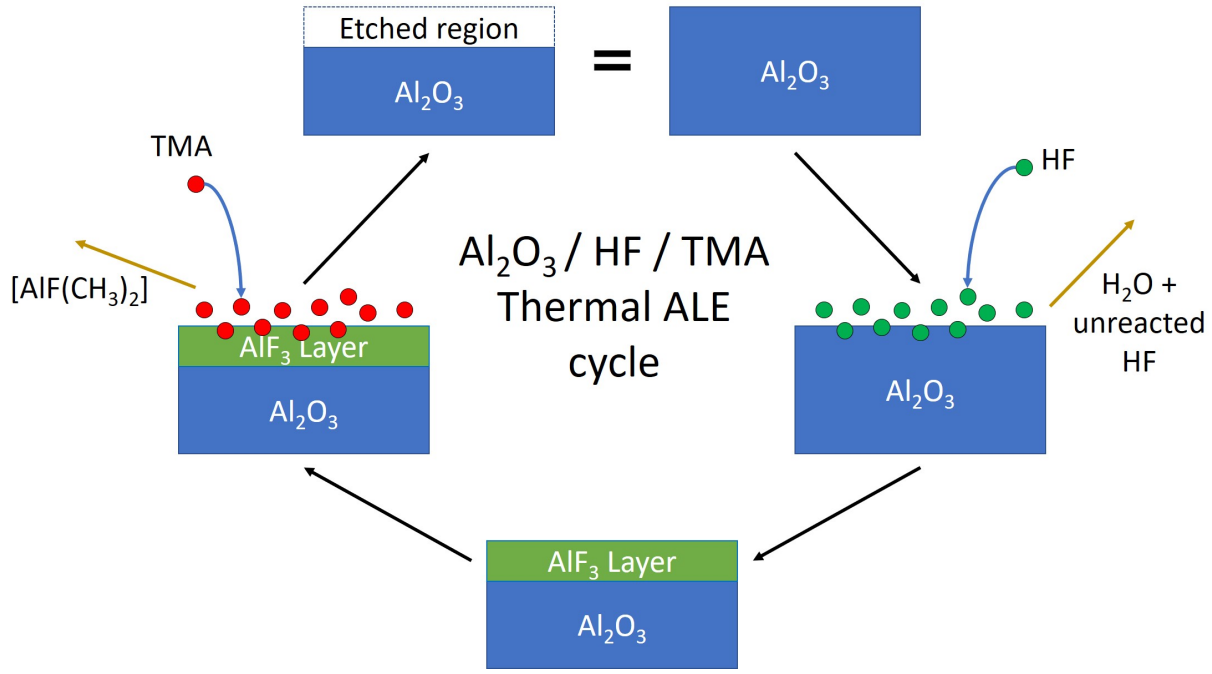


Figure 1.4: Al_2O_3 thermal ALE Process. Fluorine-activated Al_2O_3 is etched with tetramethylaluminum (TMA).

Given the two examples above, it is apparent that the defining feature of ALE is not the single atomic layer resolution of material removal but rather the digital control of etch

thickness given by the well-defined etch-per-cycle, EPC, independent of the exposure time. From a practical point of view, it is difficult to determine exactly the number of atomic layers removed. In this sense, demanding the strict "atomic-layer" level of control is not meaningful at the present.

Whereas the preceding examples are on silicon and aluminum oxide, the most dramatic increase in material choices for integrated circuits lies in metals. Metals offer many desirable traits for microelectronics integration, such as high electrical and thermal conductivities. On the other hand, effective patterning techniques are most lacking for metals. Next, the application of copper (Cu) and nickel (Ni) in the integrated circuit fabrication process is highlighted here to motivate the development of self-limiting, highly-controllable removal processes.

Copper has been the most prevalent metal in circuitry by far since the late 1990's as the standard metal for interconnects. While it boasts superior conductivity, it is at the same time notorious for migrating into the silicon substrate.[39] Additionally, the industry had been unable to develop an effective low-temperature gas phase etching process to pattern Cu. This resulted in the Damascene and chemical mechanical polishing (CMP) process, where the interlayer dielectrics are patterned lithographically, after which Cu is plated to fill the void. This clever trick avoided the need to pattern Cu,[51] but the challenges of high-aspect ratio trench filling uniformity[178] and CMP-induced cracking[63] still arose. Despite advances in Cu etching techniques at lower temperatures since then,[54,145] the Damascene and CMP process remained the industry standard. In recent years, 13 to 15 levels of metal layers are typically required to support technology nodes beyond 10 nm.[130] figure 1.5 shows a high-performance dense interconnect that meets computing needs and reliability requirements.[142] This structure clearly places a lot of constraint on the fragile low-k material, necessitating the development of a low-temperature gas phase process to etch Cu. The most viable strategy is a chemically controlled ALE process that can effectively pattern copper at designated regions without a grain size effect. A high selectivity to the barrier layer at the bottom would also be required.

The introduction of extreme ultraviolet (EUV) lithography technology represents a dras-

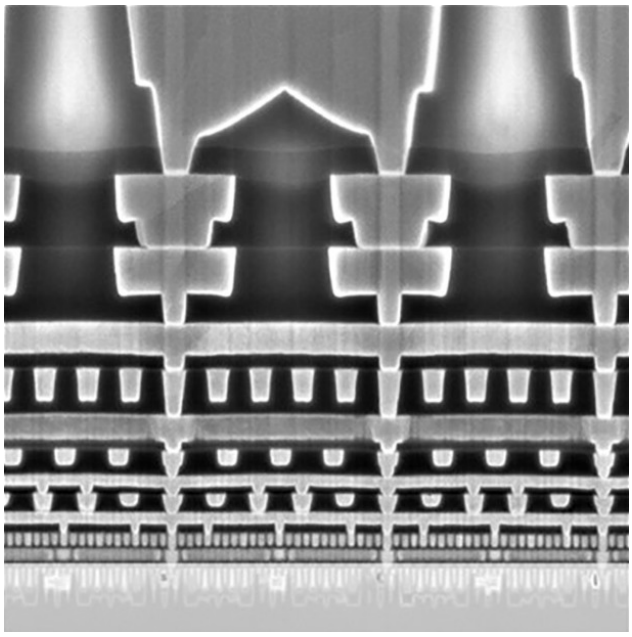


Figure 1.5: Illustration of the lowest 10 layers of metal (copper) of a 10 nm Interconnect Stack.[20]

tic transition from transmission optics to reflective optics for photolithography. The patterning of absorbers on Si/Mo multilayer-based EUV masks becomes a real challenge. It was thought that a thick layer of absorbers (e.g., 60 nm Ta) would be necessary to ensure the complete absorption of EUV in designated regions on the mask.[75] On the other hand, reducing the absorbing layer thickness generally improves the pattern transfer fidelity as the mask shadowing effect becomes less perceivable. The improvement is especially noticeable for extreme dimension features where the feature sizes are comparable to the errors introduced by shadowing. A thinner absorbing layer would require a metal that has an index of refraction (n) close to 1 and a large extinction coefficient (k). It has been measured experimentally[114] and simulated computationally[134] that Ni is such a candidate material. A thin Ni layer (30 nm) is found to be as effective as the 60 nm Ta. However, nickel integration is prevented by the challenges involved in patterning this inert metal selectively.

The copper interconnects and nickel photomask absorption layers are just two out of the many urgent applications of nanometer-level metal patterning techniques. However, in sharp contrast to the rapid rise in application needs, metal patterning on the sub-micron scale

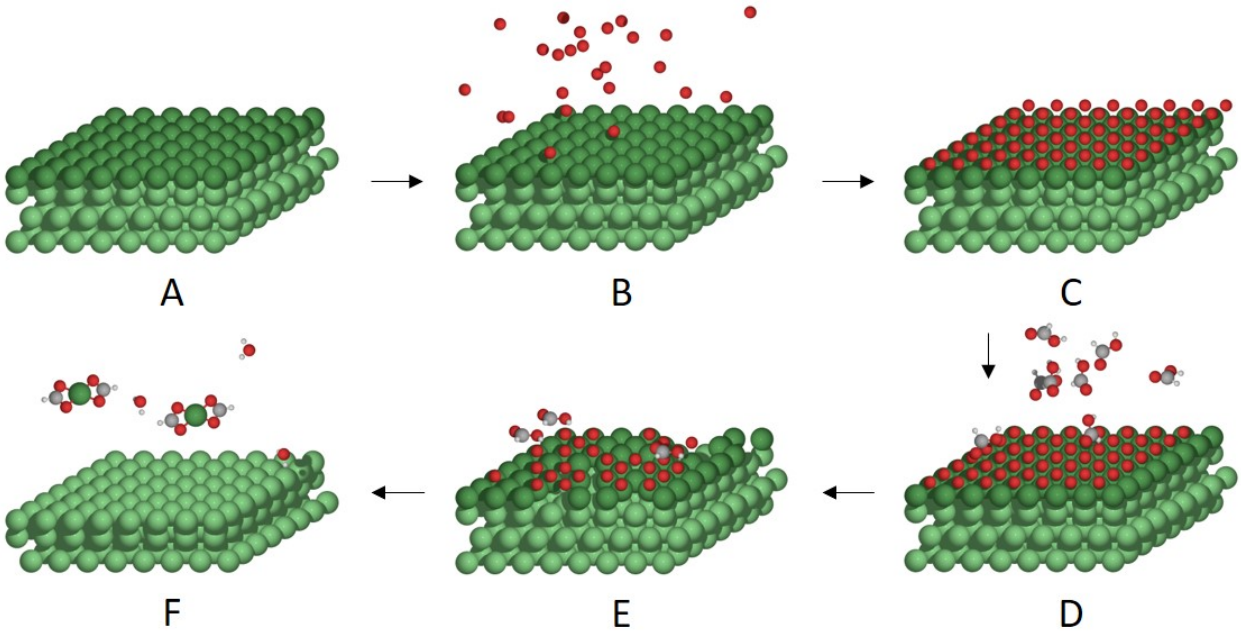


Figure 1.6: An idealized atomic layer etching (ALE) process, depicted here for Ni/O/-formic acid plasma-thermal ALE. The metal surface (A) is first subjected to a plasma ion bombardment (B) until saturation (C). An etchant molecule is then introduced (D) after the plasma has been switched off. The etchants adsorb dissociatively on the activated surface (E), etching the top layers by the formation of volatile metal-complexes. The reaction stops when all the modifier atoms are consumed, exposing the metal surface again (F). The idealized scheme here removes exactly one atomic layer (dark green), but the actual etch-per-cycle (EPC) may differ.

remains largely unsolved, as many of these metals are etch-resistant. Nickel, for instance, has long been used as an anti-corrosion additive in various alloys. Despite a few reports of controllable etching processes for metals, a systematic understanding of the underlying chemistry is still lacking.

Etching on metals is conventionally done in the solution phase byusing harsh chemicals. This is still used, for instance, in printing circuit boards. For the nanoscale patterning, dry etching involving high vacuum chambers and plasmas is now routinely used on silicon substrates. However, the techniques discussed in figure 1.2 isare not directly usable for metals, because (1) the chemical inertness of the metals requires, for dry etching, a high -energy ion source to getachieve perceivable removal, often resulting in poor selectivity, and (2) metal leaves the surface in the form of unstable, high -energy species that readily redeposit on the chamber walls and other parts of the wafer, among other reasons.

The two ALE examples given before use either plasma or thermal reactions for both modification and removal steps. In fact, these conditions can and should be combined for metals, for the reasons discussed below. To avoid metal redeposition, the removal step is best done similar to that in thermal ALE, so that metal atoms are stabilized by organic ligands. If a directional etching profile is desired, the activation step must be plasma-assisted. Plasma can also act as an energy source in the activation step. An extra source of energy is especially important for noble metals (e.g., Cu). These fundamental considerations suggest that a thermal/plasma mixed process is ideal for plasma-thermal ALE. The general features of the proposed plasma-thermal ALE process are described in figure 1.6. Proof-of-concept experimental studies have shown that this combined plasma-thermal ALE is able to etch metals (Ni and Cu) with directionality, specificity, and selectivity.

At the center of the plasma-thermal ALE process is the plasma activation of the metal substrate. Plasma is formed when the electrons in gas phase atoms/molecules are excited. The molecules can dissociate and/or ionize, creating a dilute gas of energetic free electrons and excited species. By increasing the energy of the reactant, activation energy E_a is decreased, and a normally endothermic reaction can be made exothermic since the final product on the surface is not excited. The working principle of plasma-assisted reactions is illustrated

in figure 1.7.

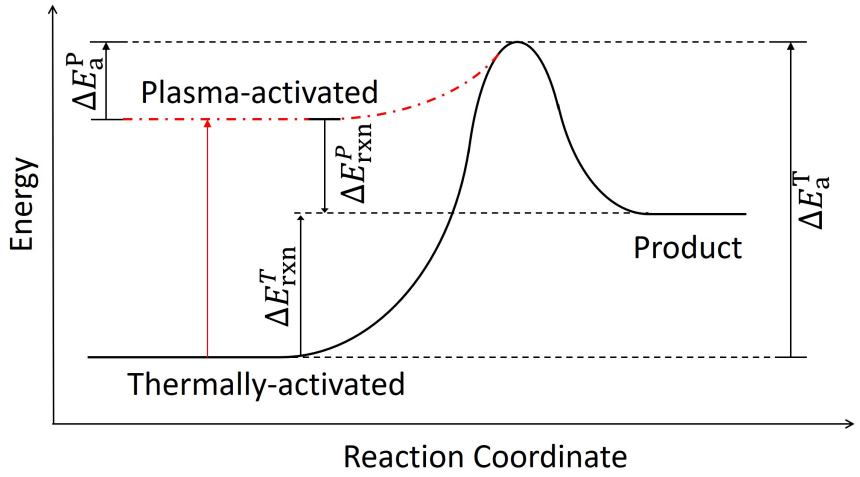


Figure 1.7: Energetics of plasma-enhanced chemistry. A thermal reaction with a positive reaction energy ΔE^T_{rxn} (endothermic) is converted to a plasma reaction with a negative ΔE^P_{rxn}. The large thermal activation energy barrier ΔE^T_a is decreased to ΔE^P_a in the plasma process, resulting in a much higher reaction rate.

Two types of plasma exist. Depending on whether the heavy species (e.g. atoms, molecules) are in thermal equilibrium with the electrons or not, the plasma is termed thermal plasma and cold plasma, respectively. A thermal plasma typically requires thousands of degrees to ionize all the species, hence cannot be used for materials processing. At the temperatures processes can operate at, electric fields can selectively transfer energy to the ions and electrons, creating a cold plasma. For our purpose, a plasma is an inter-penetrating fluid of free electrons, radicals, excited states of molecules, molecular fragments, atoms, and ions. The bulk of the plasma is charge-neutral, but a thin, negatively charged sheath is maintained at the boundary of the plasma. The sheath keeps the electrons inside the bulk plasma but accelerates the ions toward the surface. The accurate description of the chemical composition, including electrons, ground state molecules, and various dissociated and excited state molecules, is difficult[8,55] due to the inherent complexity of the reaction network and the disparate time and length scales. Fortunately, the problem can be greatly simplified by the fact that 1) ions are accelerated by the sheath to impinge nearly vertically on the surface

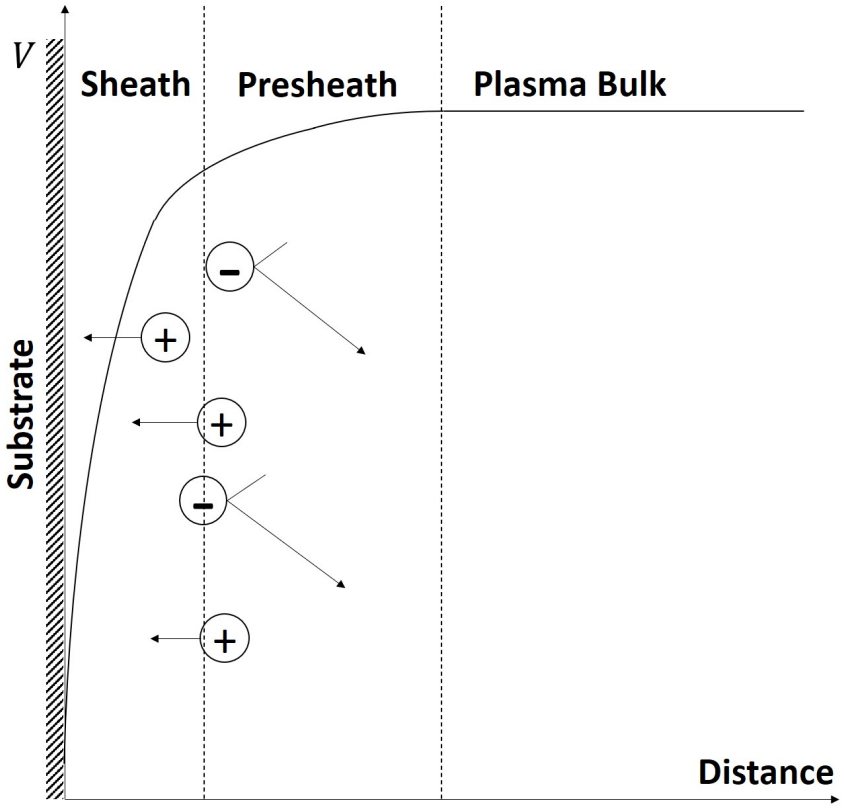


Figure 1.8: The sheath near a plasma-surface interface. The sheath creates an electric field pointing perpendicularly into the surface. This field accelerates cations toward the surface, enabling directional etching, but repels anions and electrons back into the bulk plasma to maintain the quasi-neutral condition.

with negligible speed in the lateral direction, and 2) the secondary electrons from the surface neutralize most of the ions before they reach the surface, so that ions and excited states may be approximated by fast ground state molecules. These aspects are discussed in detail in chapter 2.

The recent development of plasma-thermal ALE (Atomic Layer Etching)[161] has expanded the possibilities to include metals. However, to design and realize ALE for a wide range of metals, a thorough understanding of the underlying chemical mechanisms is essential. The validation and optimization of candidate chemistries and processes can be costly and labor-intensive if a trial-and-error approach is adopted, as reaction mechanisms depend on target materials, modifiers, and etchants. Moreover, in many cases, probing the etching chemistry experimentally is difficult due to the low concentration of gas phase products, which hinders the search for optimal chemistry and process conditions.

Computational chemistry techniques have significantly improved in accuracy over the years, providing precise estimations of thermochemical data.[118] These techniques have the potential to assist in the design and improvement of ALE chemistries, as it is easier to build and test various reactions on a computer than to perform them experimentally, provided the appropriate theoretical tools are available. As part of the eventual goal of developing a "virtual reactor" for ALE, this project serves as a proof-of-concept study using the best currently available computational tools in surface chemistry. The project can be roughly divided into thermodynamics and kinetics components.

In the thermodynamics part, the main objective is to screen chemistries (modifier and etchant pairs) for a given substrate. This means that theoretical thermochemical predictions must yield favorable energetics for a process that has already been demonstrated feasible in the laboratory, and vice versa. The energetics criteria primarily apply to the removal step. The reactivity of plasma is controllable to a great extent, so the post-plasma surface is considered an input to this model rather than a prediction output of the thermodynamics part. To facilitate the screening process, some potentially oversimplifying assumptions are necessary. However, these assumptions can be adjusted later when more information about the surface is available, without significantly affecting the model itself.

However, it is important to note that thermodynamics can only provide a lower bound for process feasibility. If a process is not thermodynamically feasible, it would not occur. On the other hand, a thermodynamically feasible chemistry might still be kinetically limited. If this is the case for the removal step, it is necessary to sample a sufficient number of low-lying pathways to obtain the required kinetic behavior. The challenges associated with this aspect are highlighted below, and an initial attempt at addressing them is provided.

The difficulty in etching energetics simulation arises from the uncertainty surrounding the post-plasma structure, which requires a realistic model of plasma-surface interaction. Simulations of plasma processing are a multiscale problem and are typically divided into four categories. At the reactor scale, Monte Carlo techniques are used to solve continuum equations for flow, temperature, and electromagnetic fields. The solutions provide information about species concentrations over time and space,[50] allowing predictions of wafer uniformity, among other factors. Increasingly, reactor scale simulations are being coupled with sheath scale simulations, such as the hybrid plasma equipment model (HPEM).[40] Sheath models aim to characterize the ion energy distribution (IED) function, a critical boundary condition for simulations at smaller scales. Feature simulations focus on the evolution of feature profiles, considering information about the flux, IED, surface reaction network, and surface diffusivity of adsorbed species.[91,101] At the smallest scale lies atomic simulations, which explicitly follow the trajectories of each atom in the process, using forces obtained from first principles or empirical interaction potentials. As devices become smaller, understanding the atomic scale etching behavior becomes increasingly important for process design. However, atomic scale simulations depend on information from larger scales, and a truly predictive multiscale model integrating all levels is still not available.

At the atomistic scale, there are two main categories of methods for treating plasma-surface interactions: kinetic Monte Carlo (KMC) and molecular dynamics (MD).[20,57,99,133] Both methods are typically trained and validated against ab initio calculations, with density functional theory (DFT) being the most common approach. In KMC, a library of events and their probabilities is required to propagate the trajectories. The main advantage is the ability to access long time scales, but this correspondence and the trajectory itself rely on the

accuracy of the library, which can be poorly transferable. Within the general MD paradigm, methods can be classified into empirical, semi-empirical, and ab initio, in order of increasing accuracy, transferability, and cost. Empirical MD methods are the fastest but often inaccurate and not transferable. DFT ab initio methods are accurate but computationally expensive, making dynamical calculations impractical for systems with more than approximately 100 atoms and time scales of a few picoseconds. Semi-empirical methods (e.g. DFTB) are typically two orders of magnitude faster than DFT, with acceptable transferability, but at the cost of sacrificing some accuracy.[33,71]

The modeling of plasma activation is complicated by the fact that the substrate is exposed to electrons, neutrals, radicals, and radiation in the plasma. While the electrons, radiation, and excited species are generally not treated in current models, describing the heavy ions alone presents a significant challenge. The kinetic energies of these ions, typically ranging from 10 eV–20 eV, far exceed the thermal energies associated with Brownian motion at practical process temperatures. Consequently, the distances between ions and substrate atoms cover larger ranges than those involved in typical thermal reactions, making interaction potentials (MD)/probabilities (MC) developed for thermal chemical reactions unusable. As a result, existing work often relies on specifically trained potentials.

For instance, the Stellinger-Weber potential is commonly employed to treat Si/Cl interactions,[26,69] while the Tersoff-Brenner potential is another popular choice.[84] In the case of metallic systems, the Embedded Atom Method (EAM) is often used,[13] while inert ions (e.g., Ar+) are typically treated with purely repulsive Moliere potentials.[31] Reactive force fields (e.g., ReaxFF) are also widely recognized for their unique ability to reproduce bond breaking/formation events. These force fields use bond-order formulations fitted to quantum chemistry data, yielding good results in related projects.[88,120,131,165]

Traditional empirical potentials are fundamentally limited in accuracy because their analytical forms are approximate. To address this limitation, artificial neural networks (ANNs) have been proposed as a potential solution since they can be fitted to complex functions.[68] This approach has been successfully applied to the simulation of complex material systems.[123,154] Another development in this area is the delta neural network potential, which

aims to improve existing empirical potentials by appending an additional NN-generated energy term, effectively fitted to the errors of the original force field.[148] This approach can be easier to implement, as much of the complexity is already treated in the empirical force field.

However, it is essential to note that in classical potentials, physics-inspired functional forms ensure that the short-distance, exponential growth of the kinetic energy is described correctly, at least qualitatively. In contrast, machine-learning potentials do not inherently provide such guarantees, and special care must be taken to obtain stable and accurate dynamics, as explored in this dissertation.

Besides the availability of accurate reactive force fields, the notorious time scale problem prevents MD simulations from reaching realistic time scales comparable to laboratory processes: accurate dynamics require discrete time steps of femtoseconds, yet many important processes (e.g. diffusion) happen on much slower timescales, and laboratory processing is in seconds or minutes. As a result, plasma impact simulations often rely on an artificially high ion incidence rate to reduce the computational cost, sacrificing accuracy and correspondence to physical time.

To overcome this problem, a class of accelerated molecular dynamics techniques have been developed.[122] These techniques can be classified into parallel replica MD, temperature accelerated MD, and hyperdynamics, of which only hyperdynamics promises to bridge the time scale gap of approximately nine orders of magnitude. Hyperdynamics adds a bias potential to the true potential energy surface in an effort to uniformly lower the transition barriers, thus accelerating the simulation of slow events (e.g. bond breaking).[41,42] While the construction of such a bias potential is far from trivial, recently a class of methods has been developed to construct bias potentials on-the-fly, eliminating the need for expert knowledge of the system on the user's end.[103,127]

On the other end, kinetic Monte Carlo methods bypass the timescale problem by converting the trajectory to a chain of discrete system states and pre-defined transition rates among them. Provided such barriers are accurate and the event library is complete, the KMC method allows the system trajectory to be mapped out very efficiently. However, KMC

is not suitable for discovery of new processes since the event library must be constructed beforehand.[72] It also does not allow the calculation of free energy, unlike MD sampling.

While the mechanisms of thermal ALE were recently reviewed,[155] the focus was mainly on high-level, abstract reaction routes (e.g., fluorination and ligand exchange). Studies involving explicit surface-etchant interaction models are limited, and they often yield qualitative and explanatory results rather than quantitative predictions. For example, some studies examined the reaction of hexafluoroacetylacetone (hfacH) with nickel and nickel oxide surfaces.[150–152] On the pristine nickel surface, the etchant was found to fragment, consistent with experimental observations. On the oxide surface, deprotonation was observed, leading to strong bonding between the nickel atom and carbonyls in the etchant molecule, forming surface hydroxyls. Desorption of the product was found to be difficult due to high electronic energies. However, the surface models used in these studies were relatively simple, limiting the scope of the results.

There are only a few examples of DFT modeling for the thermal ALE mechanisms of other metals. For instance, the reaction of hfacH and chlorine with a cobalt surface[139,144] and the reaction of acetylacetone (acacH) with iron.[167] These studies compared the energy cost of extracting metal atoms from an adatom site or from the terrace. Still, fewer studies provided kinetic information for the etching step. Some recent works calculated the kinetic energy barriers for specific steps in the thermal ALE process of cobalt[176] and studied the simpler $Si/Cl_2/Ar^+$ ALE process using time-dependent DFT (TDDFT) due to the use of photons as the desorption energy source.[177]

In contrast, the activation step has received more attention in previous studies.[109,117,139,155,157] However, it's worth noting that many of these mechanisms are educated guesses with limited experimental or computational support.[137] For instance, the activation step of thermal ALE involving HF on alumina,[119] HfO_2, and ZrO_2[157] were studied using ab-initio methods with DFT calculations of reaction pathways for HF dissociative adsorption and H_2O formation.

The apparent lack of attention on the removal step mechanism is attributed to its complexity and the uncertainty surrounding the atomistic structure of the post-activation sur-

face. In the thermal removal step, the kinetic energy from ions that could break any chemical bond in the modification step is no longer available. The limited thermal energy means that low-lying pathways must be known to accurately predict the kinetics of etching. If the macroscopically observed self-limited reaction is kinetically rather than thermodynamically limited,[60,102] relying solely on thermodynamics becomes inadequate. Additionally, the changing surface condition (depletion of the modifier and thinning of the activation layer) suggests that preference for branching reactions may be time-dependent, even if the cycle-averaged phenomena is cycle-independent. Given the uncertainty surrounding the atomistic structure of the post-activation metal surface, a full exploration of all pathways appears unwarranted.

The present dissertation aims to address the challenges discussed by developing a set of computational tools. In chapter 3, a thermodynamic model is utilized to calculate an etching free energy descriptor, which enables the screening of candidate chemistries using first principles simulation data. The accuracy and descriptive power of these models rely significantly on the chosen post-plasma surface model. Therefore, in chapter 5, an outline for the development of a machine learning-based plasma-surface interaction potential is given. This potential allows for more accurate and efficient simulations of plasma-surface interactions.

In chapter 6, the developed potential is applied in molecular dynamics simulations, which provide insights into process tuning directions for achieving a truly self-limiting plasma oxidation. The simulations predict the formation of an amorphous oxide surface. Overall, the combination of thermodynamic modeling, machine learning-based potential development, and molecular dynamics simulations offers a comprehensive approach to tackle the challenges in plasma-thermal ALE and provides valuable tools for designing and optimizing metal patterning techniques at the nanoscale.

CHAPTER 2

Theoretical backgrounds

2.1 Density Functional Theory

In this chapter, the theoretical background of the computational tools used in this dissertation is provided. The primary goal is to offer an intuitive physical understanding to serve as a starting point for interested readers.

Most of the problems in this dissertation involve finding the ground state energy of a system of M atoms and N electrons (and the inability to find it efficiently), which corresponds to solving the time-independent Schrödinger equation:[37]

$$H\Psi = E\Psi \tag{2.1}$$

Here, H is the Hamiltonian operator, given by:

$$H = -\sum_{i=1}^{N}\frac{1}{2}\nabla_i^2 - \sum_{A=1}^{M}\frac{1}{2M_A}\nabla_A^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{r_{iA}} \tag{2.2}$$

$$+\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{1}{r_{ij}} + \sum_{A=1}^{M}\sum_{B>A}^{M}\frac{Z_A Z_B}{R_{AB}} \tag{2.3}$$

Except in special cases, nuclei move much slower than electrons due to the significant mass difference. Therefore, it is reasonable to assume that electronic and nuclear degrees of freedom are decoupled, and the nuclei are considered fixed, leading to the Born-Oppenheimer approximation:

$$\sum_{A=1}^{M} \frac{1}{2M_A} \nabla_A^2 = 0 \tag{2.4}$$

$$H = H_{\text{elec}} + \sum_{A=1}^{M} \sum_{B>A}^{M} \frac{Z_A Z_B}{R_{AB}} \tag{2.5}$$

Since the nuclear Hamiltonian (second term on the right) is a constant Coulombic sum and does not affect the eigenstate, the main problem lies in solving the electronic problem:

$$H_{\text{elec}} = -\sum_{i=1}^{N} \frac{1}{2} \nabla_i^2 - \sum_{i=1}^{N} \sum_{A=1}^{M} \frac{Z_A}{r_{iA}} + \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{r_{ij}} \tag{2.6}$$

$$H_{\text{elec}} \Psi_{\text{elec}} = E_{\text{elec}} \Psi_{\text{elec}} \tag{2.7}$$

In general, Ψ represents a many-body function in the spatial and spin coordinates (x, y, z, σ) of all the electrons in the system. For N electrons, this corresponds to $4N$ degrees of freedom. This is a computationally intractable problem. For instance, if each degree of freedom can take one of 10 values, storing the wavefunction would require 10^N numbers. For a system of 100 electrons, this amounts to 10^{400} or 389 TB of data if one byte is used for each number to even store the solution, let alone solving it. Clearly, further approximations are needed to tackle this problem.

There are essentially two main approximations used to tackle this problem. In wave function methods, the many-body wave function is decoupled into a construction of one-body wave functions within the Hartree-Fock approximation:

$$|\Psi\rangle = |\chi_1 \chi_2 \cdots \chi_a \chi_b \cdots \chi_N\rangle \tag{2.8}$$

Where $|\chi_i\rangle$ is a single-body wave function (orbital) dependent only on 4 (or 3 in restricted systems), and the $|\Psi\rangle$ is a Slater determinant satisfying the Pauli exclusion principle: $|\Psi\rangle$ changes sign with respect to exchange of two spin coordinates. For the problem of storing the wavefunction, we now only need 100×10^4 or 1 MB of data. Note that Hartree-Fock by itself is hardly useful for chemistry problems, and modern wave function methods rely heavily on post-Hartree-Fock corrections.

The other direction is density functional theory (DFT). In 1964, the first, existential Hohenberg-Kohn theorem stated that "the external potential V_{ext} is (to within a constant) a unique functional of $\rho(\vec{r})$; since, in turn, $V_{\text{ext}}(\vec{r})$ fixes H, we see that the full many-particle ground state is a unique functional of $\rho(\vec{r})$."[2] This implies that, in principle, knowing $\rho(\vec{r})$ is sufficient to obtain all properties of interest. The way to obtain the density is stated by the second, variational Hohenberg-Kohn theorem: the true ground state energy can obtained from the functional only when the true ground state electronic density is used. Thus, one can find the true ground state density by the variational principle if the exact functional is known.

$$E_0 = E[\rho_0] \leq E[\tilde{\rho}] \tag{2.9}$$

In principle, density functional theory is exact. However, the Hohenberg-Kohn theorem stated nothing about the form of the true functional and, except in a few simple systems for which high-level wave function calculations are possible, various approximate functionals must be used for practical calculations.

First, the exact functional $E[\rho]$ can be written as the sum of the external potential functional, the kinetic energy functional, and the electron-electron interaction:

$$E[\rho] = E_{\text{Ne}}[\rho] + T[\rho] + E_{\text{ee}}[\rho] \tag{2.10}$$

The term $E_{\text{Ne}}[\rho]$ is simply given by

$$E_{\text{Ne}}[\rho] = \int \rho(\vec{r}) V_{\text{Ne}} d\vec{r} \tag{2.11}$$

Where V_{Ne} is the coulombic potential of the nuclei. The electron-electron interaction can be decomposed into a classical Coulomb term and a non-classical term containing effects that cannot be mapped onto a classical picture.

$$E_{\text{ee}}[\rho] = \frac{1}{2} \int \frac{\rho(\vec{r}_1)\rho(\vec{r}_2)}{r_{12}} d\vec{r}_1 d\vec{r}_2 + E_{\text{ncl}}[\rho] \tag{2.12}$$

To describe the kinetic energy, Kohn and Sham re-introduced orbitals back into the theory

by writing the kinetic energy as:

$$T_{\mathrm{KS}} = -\frac{1}{2}\sum_{i}^{N} \langle \chi_i | \nabla^2 | \chi_i \rangle \tag{2.13}$$

Where Kohn-Sham orbitals χ_i are the single-electron wave functions.[3] Of course, this choice of the kinetic functional is by no means exact, and the residual error is swept into the so-called exchange-correlation functional:

$$E_{\mathrm{KS}}[\rho] = E_{\mathrm{Ne}}[\rho] + T_{\mathrm{KS}}[\rho] + \frac{1}{2}\int \frac{\rho(\vec{r}_1)\rho(\vec{r}_2)}{r_{12}} d\vec{r}_1 d\vec{r}_2 + E_{\mathrm{XC}}[\rho] \tag{2.14}$$

The Kohn-Sham wave function is the ground state for the non-interacting reference system that gives the same density as the true ground state, if the exchange-correlation functional is known exactly. However, in practice such functional must be approximated. Note that the term "exchange-correlation" neatly summarized the effects missing from the other terms, and is inspired from the exchange integral and the correlation effects (dynamical and fermi correlation). Nonetheless, since the KS orbitals do not correspond to the single-particle orbitals that constitute a manybody wavefunction, the exchange energy calculated from the KS orbitals do not cleanly correspond to the physical exchange energy. For the same reason, one cannot use post-HF methods on these Kohn-Sham orbitals to treat correlation. This lack of a "Jacob's ladder" to progressively improve the accuracy is a major drawback for DFT.

The exchange-correlation functional is the key to the success of DFT. The simplest functional is inspired by the simple case of the uniform electron gas, for which the exchange-correlation functional is known exactly. The local density approximation, LDA, is given by:

$$E_{\mathrm{xc}}^{\mathrm{LDA}}[n_\uparrow, n_\downarrow] = \int d\mathrm{r} \epsilon_{\mathrm{xc}}^{\mathrm{unif}}(n_\uparrow, n_\downarrow) n(\mathrm{r}) = \int d\mathrm{r} [\epsilon_x^{\mathrm{unif}}(n_\uparrow, n_\downarrow) + \epsilon_c^{\mathrm{unif}}(n_\uparrow, n_\downarrow)] n(\mathrm{r}) \tag{2.15}$$

This is exact only for the fictitious system of a uniform electron gas, where the exchange ϵ_x can be calculated analytically and the correlation is accurately approximated by Monte Carlo simulations and tabulated. In real systems, the exchange-correlation energy is approximated assuming the exchange and correlation energy is the same as in the uniform electron gas, taking the *local* electronic density as the reference density.[30]

Obviously, real systems do not have a uniform density. The generalized gradient approximation (GGA) uses information from the gradient of density.[30] Most approximate functionals separately design the exchange and correlations parts, as shown in equation (2.16).

$$E_{\mathrm{xc}}^{\mathrm{GGA}}[n_\uparrow, n_\downarrow] = E_{\mathrm{x}}^{\mathrm{GGA}} + E_{\mathrm{c}}^{\mathrm{GGA}} \tag{2.16}$$

$$E_{\mathrm{x}}^{\mathrm{GGA}}[n_\uparrow, n_\downarrow] = \frac{1}{2}\{E_{\mathrm{x}}[2n_\uparrow], E_{\mathrm{x}}[2n_\downarrow]\} \tag{2.17}$$

The exchange energy functional is typically given in terms of an "enhancedment factor" that depends on the dimensionless gradient t. A typical functional form is given by equation (2.18):

$$E_{\mathrm{x}}(\mathbf{r}) = \int \mathrm{d}^3 r E_{\mathrm{x}}^{\mathrm{unif}}[\rho(\mathbf{r})] \cdot f_{\mathrm{x}}(t(\mathbf{r})) \tag{2.18}$$

$$t(\mathbf{r}) = \frac{|\,\rho(\mathbf{r})|}{2k_{\mathrm{F}}(\mathbf{r})\rho(\mathbf{r})} \tag{2.19}$$

With $f_{\mathrm{x}}(t)$ designed to satisfy certain known constraints of the exact functional. The correlation functional is given by the sum of a LDA contribution and the gradient-dependent contribution H, as shown in equation (2.20).

$$E_{\mathrm{c}}[n_\uparrow, n_\downarrow] = \int d\boldsymbol{r} n[\epsilon_{\mathrm{c}}^{\mathrm{unif}}(r_s, \zeta) + H(t, \rho, \zeta)] \tag{2.20}$$

Where $\zeta = (n_\uparrow - n_\downarrow)/n$ is the dimensionless spin polarization.

Note that in contrast to the LDA functional which is grounded firmly in the physical picture of a uniform electron gas, the GGA functionals have considerable freedom in their design. Various forms / parametrizations of the enhancement factor f_{x} in the exchange and the gradient-dependent contribution H to the correlation have been proposed, relying either on empirical fit or various known constraints on the exact functional. Therefore, there are many GGA functionals. In general, the choice of the exchange-correlation functional is problem-dependent and the results must be carefully benchmarked.

The work presented in this thesis uses the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional.[36] This functional has been shown to yield good accuracy for studying the adsorption on metal surfaces, as well as generally good performance in the solid

state.[107,143] The functional is given by:

$$f_{\mathrm{x}}(s) = 1 + \kappa - \kappa/\left(1 + \mu s^2/\kappa\right); \tag{2.21}$$

$$H = \left(e^2/a_0\right)\gamma\phi^3 \ln\left\{1 + \frac{\beta}{\gamma}t^2\left[\frac{1 + At^2}{1 + At^2 + A^2t^4}\right]\right\} \tag{2.22}$$

$$A = \frac{\beta}{\gamma}\left[\exp\left\{-\epsilon_{\mathrm{C}}^{\mathrm{unif}}/\left(\gamma\phi^3 e^2/a_0\right)\right\} - 1\right]^{-1}; \tag{2.23}$$

Here, spin polarization is accounted for in the dimensionless gradient $s = \frac{|\nabla n|}{2\phi k_s n}$. r_s is the density given in Seitz radius as $n = 3/4\pi r_s^3$, $k_s = \sqrt{4k_F/\pi a_0}$ is the Thomas-Fermi screening wave number (a_0 is Bohr radius). $\phi(\zeta) = \left|(1+\zeta)^{2/3} + (1-\zeta)^{2/3}\right|$ is a spin-scaling factor. The remaining symbols $\kappa = 0.804, \mu = 0.21951, \beta \approx 0.066725, \gamma \approx 0.031091$ are all physically-derived constants.

2.2 The self-consistent field solution of Kohn-Sham DFT

With the functionals defined, the problem translates to finding a self-consistent solution of the Kohn-Sham eigenvalue problem:

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + v_{\mathrm{eff}}(\vec{r})\right)\psi_i(\vec{r}) = \varepsilon_i\psi_i(\vec{r}) \tag{2.24}$$

The effective one-particle potential V_{eff} is a functional of the density, as shown in equation (2.25):

$$v_{\mathrm{eff}}(\mathbf{r}) = v_{\mathrm{ext}}(\mathbf{r}) + e^2\int\frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}\,d\mathbf{r}' + \frac{\delta E_{\mathrm{xc}}[\rho]}{\delta\rho(\mathbf{r})} \tag{2.25}$$

The density is simply the sum of squares of the orbitals (f_i is the occupation):

$$\rho(\vec{r}) = \sum_i^N |\psi_i(\vec{r})|^2 f_i \tag{2.26}$$

The self-consistent field method iteratively update the KS orbitals ψ_i, typically expanded in terms of the analytical basis functions:

$$|\psi_i\rangle = \sum_\alpha c_{i,\alpha}\,|\phi_{i,\alpha}\rangle \tag{2.27}$$

Assuming a complete basis set is used, at each step of the SCF iteration scheme, the problem is translated to a linear eigenvalue problem over the Hilbert space spanned by the basis

functions $\{\phi\}$. Upon successful solution of this problem, V_{eff} is updated by the new density, which forms a new eigenvalue problem. The self-consistent solution is obtained when the eigenvalues do not change between iterations.

Note that in contrast to Hartree-Fock, the Kohn-Sham equation does not involve the manybody wave-function. The manybody effect is treated as a mean-field approximation of the density, where every other orbital contributes only through the one-body integrals $|\psi_i(r)|^2 dr$. In other words, there are no costly construction of overlap and exchange integrals to evaluate the effective operator, as in Hartree-Fock theory. This is the fundamental reason for the computational efficiency of DFT. With modern exchange-correlation functionals, an accuracy similar to the high-level post-HF methods is routinely achieved at a tiny fraction of the cost.

2.3 The PAW method

As explained above, the efficient solution of the self-consistent field (SCF) problem requires a basis set. In solid state system, a natural choice of the basis set is the Bloch functions:

$$\phi(\vec{r}) = \exp\Big(i\vec{k}\cdot\vec{r}\Big)u(\vec{r}) \tag{2.28}$$

where $u(r)$ is a periodic function periodic for any vector T commensurate with the unit cell:

$$u(\vec{r}+\vec{T}) = u(\vec{r}) \tag{2.29}$$

Projecting the electronic wavefunction ψ onto such plane waves is efficient with fast fourier transforms (FFT) when ψ varies slowly in space. This is true for regions far away from the nuclei. However, for regions close to the nuclei, the wave function varies rapidly with respect to space. An accurate projection thus requires many high-frequency (large $|\vec{k}|$) plane waves, making it computationally very inefficient. The pseudopotential approach sought to address the efficiency problem by replacing core electrons with an effective potential. In this approach, only the slow-varying valence electron wave functions are represented using a plane wave basis, allowing a much reduced basis set to be used.

The key idea behind the PAW method is to introduce a linear transformation operator T, known as the augmentation operator, that transforms the pseudowave $\left|\widetilde{\psi}_{n\mathbf{k}}\right\rangle$ to the all-electron wavefunction $|\psi_{n\mathbf{k}}\rangle$. The pseudowave, similar to the pseudopotential method, varies smoothly near the core region, while the transformed all-electron waves preserve the fast-varying nodal features but need not be represented during calculation.[7,19,28] The operator $T_R = \sum_i \left(|\phi_i\rangle - \left|\widetilde{\phi}_i\right\rangle\right) \langle\tilde{p}_i|$ is only non-zero in the region within distance R from the nuclei, ensuring that the total electron density and the valence electron density obtained from the all-electron wavefunctions match the corresponding densities obtained from the pseudized wavefunctions:

$$|\psi_{n\mathbf{k}}\rangle = \left|\widetilde{\psi}_{n\mathbf{k}}\right\rangle + \sum_i \left(|\phi_i\rangle - \left|\widetilde{\phi}_i\right\rangle\right) \left\langle\tilde{p}_i\middle|\widetilde{\psi}_{n\mathbf{k}}\right\rangle \tag{2.30}$$

$$T = 1 + \sum_i \left(|\phi_i\rangle - \left|\widetilde{\phi}_i\right\rangle\right) \langle\tilde{p}_i| \tag{2.31}$$

The pseudowaves $\left|\widetilde{\psi}_{n\mathbf{k}}\right\rangle$ are expanded in the plane wave basis (equation (2.28)). The all-electron (AE) partial waves ϕ_α are the solutions to the radial schrodinger equation for the non-spin polarized atom. By construction the pseudo partial waves $\tilde{\phi}_\alpha$ are equal to ϕ_α outside the core region.

The PAW method provides several advantages over pseudopotential methods.,[27] Firstly, it brings the core-electrons back into the Kohn-Sham orbitals, allowing for electron-electron interactions between the core and the valence to be evaluated. This is important for systems where the core electrons play a significant role in the electronic structure. Secondly, this method allows the reconstruction of all-electron waves from pseudowaves, making it more accurate for total energy, forces, and response functions. Finally, these added advantages lead to improved convergence with respect to plane wave cutoff and faster calculations.

2.4 Ab-initio Thermodynamics

While DFT operates at 0 K to reach the electronic ground state, chemistry occurs at finite temperatures. Therefore, it is desirable to model the energetics accounting for thermodynamics at various temperatures, pressures, compositions, *etc.*, since it is the free energy G, not the electronic potential energy, that governs reaction spontaneity. Statistical mechanical ensembles provide the necessary link between electronic structure and thermodynamics. Specifically, the effects of zero-point energy, vibrations, rotation, and translation need to be accounted for. The free energy is given by:

$$G = H - TS \tag{2.32}$$

For an ideal gas (all gaseous systems considered here are dilute, therefore satisfying the ideal gas assumptions). The enthalpy H is given by:

$$H(T) = E_{\text{elec}} + E_{\text{ZPE}} + \int_0^{\text{T}} C_P \, \mathrm{d}T \tag{2.33}$$

The zero-point energy is simply $\frac{1}{2}\sum_i \hbar\omega_i$. The heat capacity is evaluated from the canonical partition function.

$$C_P = k_{\text{B}} + C_{V,\text{trans}} + C_{V,\text{rot}} + C_{V,\text{vib}} + C_{V,\text{elec}} \tag{2.34}$$

In equation (2.34), the equations to evaluate the terms are well documented and can be found in standard stastical mechanics textbooks.[62] For computational evaluation, $C_{V,\text{trans}}$ requires the molar mass of the molecule, $C_{V,\text{rot}}$ requires the rotational inertia, and $C_{V,\text{vib}}$ requires the vibrational eigenmodes. The electronic contribution $C_{V,\text{elec}}$ is not important at temperatures considered here. A similar observation is made for the relevant terms in the entropy contribution to free energy:

$$S(T,P) = S(T,P^\circ) - k_{\text{B}} \ln \frac{P}{P^\circ} \tag{2.35}$$

$$= S_{\text{trans}} + S_{\text{rot}} + S_{\text{elec}} + S_{\text{vib}} - k_{\text{B}} \ln \frac{P}{P^\circ} \tag{2.36}$$

Therefore, the only additional computation needed for thermochemistry is the calculation of vibrational eigenmodes (phonon modes for solids). These frequencies can be calculated using finite difference or linear response methods (density functional perturbation theory, DFPT).

2.5 Machine learning interatomic potentials

In order to perform large-scale dynamical simulations, a force field is needed since density functional theory does not scale well with respect to the number of electrons. Traditional force fields, such as empirical potentials or semi-empirical methods, are often based on predefined functional forms and parameters. In contrast, machine learning interatomic potentials (MLIPs) use highly flexible functional forms, combined with, in some cases, complete descriptors of the local atomic environment, to directly interpolate the atomic potential energy surface. These potentials are designed to accurately describe the interactions between atoms in a system.

The training process involves generating a diverse set of atomic configurations. The ground-state electronic energies are used as training labels. The trained potential interpolates between atomic coordinates in the training dataset using the provided labels. In some cases, the corresponding DFT-derived Hellmann-Feynman forces and stresses are also used for training.

While a myriad of methods exist, MLIP models generally consist of two parts. The featurization part converts the 3-dimensional atomic configurations to a feature vector, or, in kernel-based methods, a vector of kernel-transformed environments. Featurization encodes relevant information about the local atomic environment. It is important that the featurizers are translationally, rotationally, and permutationally invariant (see below). Respecting such physical symmetries typically results in much better data efficiency and improved accuracy. The state-of-the-art featurizers are summarized in a recent review..[171]

After featurization, the MLIP transforms the features into an output "atomic energy". The transform can be as simple as a linear combination of the features, or as complex as message-passing neural networks. As an intuitive rule of thumb, higher resolution in the featurization allows the use of less nonlinearity in the transform to achieve a given accuracy, and vice versa. Combining a high-resolution featurizer and highly flexible nonlinear transformation leads to state-of-the-art models. As a last step, the atomic energy contributions are aggregated (usually a summation pool) to yield the system's total energy. Forces are

obtained as the negative gradient of atomic positions, calculated by analytical expressions obtained from chain-ruling back to the featurizers and auto-differentiation through the non-linear part. Note that the "atomic energy" contributions do not have a physical meaning but merely serve as a mathematical trick to enable size-extensivity of the predicted total energy. See section 5.5 for an discussion about the potential problems.

As an example of such models, the Behler-Parinello high-dimensional neural network potential is described. Here, the local atomic environments are featurized using the so-called atom-centered symmetry functions (ACSFs), whose functional forms are given in equations (2.38) and (2.39) for the radial and angular variants.

$$G_{i,\mu}^{\mathrm{rad}} = \sum_{j\neq i}^{N_{\mathrm{atom}}\in R_{\mathrm{c}}} \mathrm{e}^{-\eta\left(R_{ij}-R_{\mathrm{s}}\right)^2} f_{\mathrm{c}}\left(R_{ij}\right) \tag{2.37}$$

$$G_{i,\mu}^{\mathrm{ang}} = 2^{1-\zeta} \sum_{i,j,k} \left(1+\lambda\cos\theta_{ijk}\right)^{\zeta} \mathrm{e}^{-\eta\left(R_{ij}^2+R_{ik}^2+R_{jk}^2\right)} f_{\mathrm{c}}\left(R_{ij}\right) \tag{2.38}$$

$$f_{\mathrm{c}}\left(R_{ik}\right) f_{\mathrm{c}}\left(R_{jk}\right) \tag{2.39}$$

Where η, ζ, λ are hyperparameters, typically chosen according to certain heuristic rules. The functions f_c are smooth cutoff functions that enforce the locality of featurization. Typically, a reasonably accurate model (energy mean error $\lessapprox 10\,\mathrm{meV/atom}$) requires around $N = 100$ such symmetry functions, resulting in a vector of N dimensions, thus the name "high-dimensional" neural network. The feature vector is then sent through a multilayer perceptron typically with a few (around 3) hidden layers. At the output node only a single energy label is produced, and the forces are back-propagated by differentiating the total energy. The length of feature vectors, the number of hidden layers, and the number of nodes in each hidden layer, together with the parameters in the symmetry functions, constitute the hyperparameters in this model and can be optimized by standard hyperparameter optimization methods (*e.g.*, cross-validation).

Note that the ACSFs include only 2-body and 3-body interactions. As shown in recent work, such schemes are limited and cannot resolve certain differences in the environment, and higher body-order terms are needed to be complete. This represents a fundamental limit to accuracy attainable with such descriptors, but practically this is often not a significant

concern as successful models were demonstrated for a wide range of chemical environments.

In state-of-the-art models, equivariant models were proposed with great success for reproducing the accuracy. These models relax the strict invariance requirement but instead propose strict rules on how features may vary with respect to *e.g.* certain symmetry operations. For instance, it is obvious that the forces acting on the atom should be rotated along with the local atomic environment. The equivariant models allow such information to propagate through the model (typically a message-passing neural network). Doing so preserves more information at intermediate stages in the nonlinear part and leads to better data efficiency and extrapolation behavior.

As with all machine learning methods, it is important to validate machine learning interatomic potentials carefully. Testing against experimental data and reference calculations is necessary to assess their accuracy and reliability for specific applications. This is particularly true when applying MLIPs to environments not encountered in their training data (*i.e.*, extrapolating).

CHAPTER 3

Thermodynamic Model of Atomic Layer Etching

3.1 Introduction

As the general aspects of ALE modeling have already been laid out in chapter 1, here, only the parts relevant for the thermodynamic study are emphasized briefly.

Atomic layer etching, derived from plasma etching, involves two self-limited steps. In the activation step, the surface's top atomic layer is transformed by modifiers, constrained by the formed surface compound. Then, in the removal step, etchant molecules eliminate the surface compound until depletion. This cycle repeats until the desired etch thickness is achieved.

While the underlying principles remain consistent, the actual processes employed vary in terms of chemistries and conditions. Specifically, the process being considered here is plasma-enabled ALE on metals. Metal is a leading choice for integrated circuits due to its desirable traits like high electrical and thermal conductivity. However, using metals in microelectronics requires new metal patterning techniques. Conventional metal etching involves harsh chemical solutions, resulting in isotropic etching due to random molecular motion. To achieve anisotropic etching, modern techniques utilize the directional behavior of ions in the plasma sheath. While plasma etching has been successful with silicon-based materials, it poses challenges for metals. Metal species form unstable, high-energy compounds during plasma etching, leading to redeposition on chamber walls and the wafer.

A recent innovation combines plasma activation and thermal removal to achieve directional etching of Ni and Cu.[159,160] This process, termed plasma-thermal ALE, utilizes an oxygen plasma for surface activation and organic etchants for removal, eliminating the need

for an additional energy source and preventing redeposition. The plasma-driven activation imparts directionality, resulting in an anisotropic etching profile.[159,160]

A recent review by Fischer[166] provides an excellent overview of experimental and theoretical work on thermal atomic layer etching (ALE). Another review in 2020 by George et al.[155] discusses various strategies for implementing thermal ALE on different substrates. Current thermal ALE processes for metal substrates involve oxidizing the metals with oxygen (O2 and/or O3) or halogens, followed by the introduction of diketone molecules to form volatile metal complexes. Several studies have proposed mechanisms using a combination of experimental and computational approaches,[144,151,157] allowing for the classification of existing ALE processes. However, in-situ characterization challenges and low product pressures limit the confirmation and atomistic understanding of these mechanisms.

To address these limitations, a computational thermodynamic model known as the Natarajan-Elliott analysis (N-E analysis) has been developed.[172] This model focuses on the temperature dependence of the thermal ALE fluorination step and incorporates surface models at different coverages, improving upon previous work. However, the N-E analysis does not consider the effects of surface morphology due to arbitrarily chosen crystal terminations and coverages. Additionally, it primarily focuses on the activation step, whereas plasma-thermal ALE relies on energetic ions for activation, with the substrate temperature playing a secondary role in determining the post-activation surface.

Surface morphology is highly relevant in plasma-thermal atomic layer etching (ALE). The current setup lacks a self-limited oxidation reaction, as oxides with thicknesses on the order of ~ 5 nanometers are observed. Modifying the plasma's modifier or ion energy could potentially create a self-limited process that offers better control. These drawbacks highlight the need for a computational model to quickly determine the feasibility of proposed processes under new conditions and/or chemistries. Ideally, given the specifications of the substrate, modifier, etchant, and relevant reaction conditions (temperature, pressure, and ion energies), one should be able to predict, on a thermodynamic basis, whether an ALE process would be possible.

This work aims to provide a thermodynamic prediction for the removal step of plasma-thermal ALE processes. The core of the model is the layer model energy expression, which allows the calculation of the removal step energy change from quantum chemical calculations on a slab structure. The activated surfaces, after interaction with the modifiers, are sampled using the "greedy" algorithm for adsorption configuration search. The etching products are determined computationally, utilizing literature data as input structures whenever available. The removal step energies on different terminations are combined via a weighted sum, with weights obtained through equilibrium crystal shape construction. This combined approach yields a single-valued prediction for the removal step energies based on post-plasma surface structure, modifier, and etchant. Thus, the model establishes a link between post-activation surface morphology and removal step thermodynamics, which can be used to evaluate and screen potential plasma-thermal ALE processes.

The systems chosen to test the models involve two metal substrates (Cu and Ni), two modifiers (O and N), and two etchants (formamidine (HFAmd) and formic acid (HFA)). Ni, O, and formic acid are selected for computational investigation due to their demonstrated process viability. Additional chemistries involving Cu, N, and HFAmd are included to identify useful trends. Cu is the most commonly used metal in circuitry, particularly as an interconnect material. As feature sizes shrink, the existing process that combines additive patterning (damascene process) with chemical-mechanical polishing (CMP), developed to overcome the inability to etch Cu, becomes too harsh for the bottom layers of the interconnect. Therefore, new "gentle" etching approaches must be developed for this specific application. N is chosen because a nitridation-based activation process allows for easier *ex situ* surface characterization compared to oxidation-based processes. The inclusion of formamidine is inspired by the success of the amidinate family of ligands in atomic layer deposition and the fact that the removal step in plasma-thermal ALE is the reverse reaction.

It is important to note that the results presented here solely consider thermodynamics and do not account for reaction barriers. While kinetic barriers are not expected to be significant in the activation step, they may play a role in the removal step, particularly in the formation of hydrides and the desorption of formed complexes.[150,152] Existing data suggests

that desorption of the metal-organic complex is the rate-limiting step, often requiring high temperatures to overcome the barrier.[166] However, the magnitude of these barriers is likely dependent on surface morphology.[151] The post-plasma surface morphology remains unclear, and it should be noted that the surface structure employed in this study may not precisely correspond to the actual post-plasma surface. Therefore, the impact of ion kinetic energy represents a missing link that is explored in chapters 5 and 6.

3.2 Methods

The reactions studied are listed in reaction {3.1a} - reaction {3.3b}. HL refers to a generic protonated etchant since formic acid and formamidine etchings have the same stoichiometry. The symbol | indicates surface species, as in the Ni|Ni species for the Ni etching reaction (reaction {3.1a}). The right part (|Ni) represents the surface atomic layer that is etched, and the left part (Ni|) represents the subsurface part of the slab that is exposed after one full cycle. For the discussion below, the Ni/O/HFA system is used as our working example. The etching reactions on the pristine and activated surfaces are given respectively by reaction {3.1a} and reaction {3.2b}, where the activated surface is modeled by a NiO slab. The "bulk model" expressions for the Gibbs free energy change for the surface layer removal step are given in equation (3.1) and equation (3.2), respectively. The nickel complexes have a range of nuclearities from monomer to tetramer. For simplicity, these molecules are referred to as $Ni(L)_2$. The specific oligomer used is clear from the context. Notation for Cu complexes is not ambiguous since only the dimers are used (see Page 37).

$$\mathrm{Ni|Ni} + 2\,\mathrm{HL} \longrightarrow \mathrm{Ni|} + [\mathrm{Ni(L)_2}] + \mathrm{H_2} \qquad \{3.1a\}$$

$$\mathrm{Cu|Cu} + \mathrm{HL} \longrightarrow \mathrm{Cu|} + \frac{1}{2}\,[\mathrm{Cu_2(L)_2}] + \frac{1}{2}\,\mathrm{H_2} \qquad \{3.1b\}$$

$$\mathrm{Ni_3N|}\frac{2}{3}\,\mathrm{Ni_3N} + 2\,\mathrm{HL} \xrightarrow{\mathrm{Ni(L)_2}} + \frac{2}{3}\,\mathrm{NH_3} + \mathrm{Ni_3N|Ni} \quad \{3.2a\}$$

$$\mathrm{NiO|NiO} + 2\,\mathrm{HL} \longrightarrow \mathrm{NiO|} + [\mathrm{Ni(L)_2}] + \mathrm{H_2O} \quad \{3.2b\}$$

$$\mathrm{Cu_3N|}\frac{1}{3}\,\mathrm{Cu_3N} + \mathrm{HL} \longrightarrow \mathrm{Cu_3N|} + \frac{1}{2}\,[\mathrm{Cu_2(L)_2}] + \frac{1}{3}\,\mathrm{NH_3} \quad \{3.3a\}$$

$$\mathrm{Cu_2O|}\frac{1}{2}\,\mathrm{Cu_2O} + \mathrm{HL} \longrightarrow \mathrm{Cu_2O|} + \frac{1}{2}\,[\mathrm{Cu_2(L)_2}] + \frac{1}{2}\,\mathrm{H_2O} \quad \{3.3b\}$$

The plasma-thermal ALE process is made possible by a change in the metal oxidation state during the activation step. Take our working example, metallic Ni is converted from Ni(0) to Ni(II) after activation. During the etching step, the Ni(II) oxidation state is preserved. In the absence of activation, the change of oxidation state would occur during the etching, with Ni oxidizing to Ni(II), and protons from the etchant reducing to form H_2. The same can be said of copper: Cu has an oxidation state of +1 in both the activated surface and the product complex.

This brief consideration shows that the redox potential of the metal is a key factor for the practicality of plasma-metal ALE processes, and in general any oxidation state-based metal etching processes. For both metals studied, the oxidation of the pristine surface with protons is shown to be highly unfavorable on both metals studied, a prerequisite for the self-limiting behavior observed.

$$\Delta G_{\mathrm{rm,Ni}} = [G_{[\mathrm{Ni(L)_2}]} + G_{\mathrm{H_2}} + G_{\mathrm{Ni|}}] - [G_{\mathrm{Ni|Ni}} + 2G_{\mathrm{HL}}] \quad (3.1)$$

$$\Delta G_{\mathrm{rm,NiO}} = [G_{[\mathrm{Ni(L)_2}]} + G_{\mathrm{H_2O}} + G_{\mathrm{NiO|}}] - [G_{\mathrm{NiO|NiO}} + 2G_{\mathrm{HL}}] \quad (3.2)$$

The bulk model simplifies the problem by assuming that the same pristine nickel/nickel oxide surface is exposed after one cycle. Therefore, the surface formation energies are unchanged, and surface effects cancel out. For the pristine surface, the two surface terms cancel out to give the bulk formation energy:

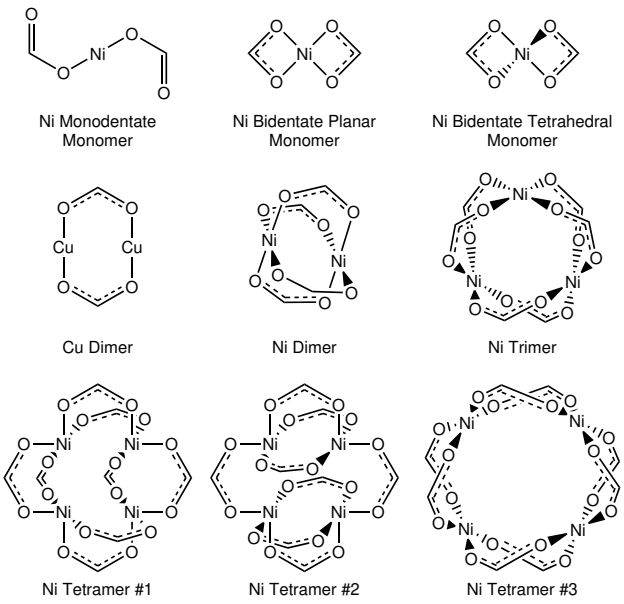


Figure 3.1: Structures of the Ni(II) and Cu(I) complexes considered. Only the formates are shown. Respective formamidinates can be found by replacing oxygen atoms with secondary amine groups. The relative stabilities are shown in figure 3.2. The dimer and tetramer structures are found to be the most stable for formate and formamidinate complexes, respectively. Oligomers are not investigated for Cu (see Page 37).

$$G_{\mathrm{Ni|}} - G_{\mathrm{Ni|Ni}} = -G_{\mathrm{Ni,bulk}}$$

The bulk model further assumes that the activated surfaces are thick enough so that they can be approximated as a bulk oxide or nitride. Under this assumption, removing one atomic layer from the activated surface is equivalent to removing one layer from the bulk oxide/nitride lattice, energetically corresponding to the Gibbs free energy of formation of bulk oxide or nitride.

$$G_{\mathrm{NiO|}} - G_{\mathrm{NiO|NiO}} = -G_{\mathrm{NiO,bulk}}$$

$$G_{\mathrm{Ni_3N|}} - G_{\mathrm{Ni_3N|Ni_3N}} = -G_{\mathrm{Ni_3N,bulk}}$$

The energy expression of the bulk model is given in equation (3.3) - equation (3.4) for etching on pristine and activated Ni systems. Note that this model avoids any calculation

of surface species, and hence can be applied easily. However, since the bulk oxide or nitride structure is assumed, this model cannot be applied to processes where only a thin surface activation layer is expected. For such processes, the layer model would be more appropriate.

$$\Delta G_{\mathrm{rm,Ni}} = G_{[\mathrm{Ni(L)_2}]} + G_{\mathrm{H_2}} - 2G_{\mathrm{HL}} - G_{\mathrm{Ni,bulk}} \tag{3.3}$$

$$\Delta G_{\mathrm{rm,NiO}} = G_{[\mathrm{Ni(L)_2}]} + G_{\mathrm{H_2O}} - 2G_{\mathrm{HL}} - G_{\mathrm{NiO,bulk}} \tag{3.4}$$

The structures and energies of the bulk crystals used are detailed in the table A.1. The molecular structures of the metal formate complexes considered are shown in figure 3.1. The corresponding formamidinate complexes are trivially found by replacing oxygen atoms with secondary amine groups. These structures are based on a combination of experimental[4–6,29,58,64,65,67,81] and simulated results.

X-ray diffraction spectroscopy on a range of copper(I) amidinates of the formula [R'NC(R)NR"Cu]$_2$ (where R' and R" = *n*-propyl, isopropyl, *n*-butyl, isobutyl, *sec*-butyl, *tert*-butyl; R = methyl, *n*-butyl) in the solid phase suggest a planar dimer structure.[64] Early mass spectroscopy suggests that the Cu formate dimerizes in the vapor phase.[4,5] Although no direct structural determination exists for the formate complex, the related copper acetate complex is shown to crystallize in the $P2_1/m$ space group as infinite sheets of 8-membered rings.[6] The high saturation pressure of the complex[5] suggests that there is little structural difference between the vapor and solid phases of the copper acetate complex. Taking clues from these, the copper(I) complexes are assumed to have a dimer structure that forms an 8-membered ring containing 2 Cu atoms.

The nickel complexes (written as Ni(L)$_2$ for simplicity, by factoring out the nuclearity) have extended covalent bonds in the solid phase; hence their gas-phase structures have been searched by constructing different structures and comparing their stabilities. The structures of the determined complexes are shown in figure 3.1.

The layer model accounts for the effects of surface termination and coverage, defined as the area density of modifier atoms. Here the stoichiometry is allowed to differ from that of the bulk model. The post-etch surfaces are assumed to terminate in the same orientation

as the pre-activation surfaces. A full-layer removal is still desirable as it avoids the explicit consideration of the post-removal surface. Therefore, the post-etch surface is assumed to be modifier-free. The energies of dimers, trimers, and tetramers of the nickel complex are normalized to the monomer formula unit for use in equation (3.3) - equation (3.4).

Suppose, for instance, that the modifier coverage after activation is less than that specified by reaction {3.2b} for the Ni/O/formic acid system. Since the reaction is constrained to proceed at the stoichiometry, such "insufficient" coverages would result in an incomplete etching of the surface layer, with Ni islands remaining at the termination of the etching process because the pristine surface cannot be etched by the etchant (see section 3.4.2). To account for this, the surface oxide is multiplied by a factor so that just enough oxygen atoms are present in the balance ($\frac{x}{y} \times y = x$ O atoms) to satisfy the stoichiometry required for the etching of x Ni atoms. This ensures that the stoichioimetry of the etching reaction is maintained, even with "insufficient" modifier coverage.

$$\frac{x}{y}\,\mathrm{Ni_s|Ni_xO_y} + 2\,x\mathrm{HL} \longrightarrow \mathrm{Ni_s|Ni_0O_0} + (\frac{x}{y} - 1)\,\mathrm{Ni_s|Ni_xO_0} + x\,\mathrm{NiL_2} + x\,\mathrm{H_2O} \qquad \{3.4\}$$

The extra $(\frac{x}{y} - 1)$ activated nickel slabs that are brought into the reaction balance only serve contribute the oxygen atoms. The unetched metal atoms correspond to the islands that remain on the surface after etching (one layer higher than the etched regions). The notations s and x refer to the structure models used, with s denoting the total number of metal atoms in the un-activate layers of the slab and x denoting the number of metal atoms per layer. Hence, $\mathrm{Ni_s|Ni_0}$ and $\mathrm{Ni_s|Ni_x}$ are slab models differing by one layer in vertical thickness. See the next paragraph for more details. The coefficients in reaction {3.4} and the factors of Gibbs free energies in equation (3.5) would depend on the specific substrate/modifier/etchant combination. The molecular terms in reaction {3.4} are obtained directly from the bulk model reaction. The reaction energy normalized to one Ni atom removed is expressed in equation (3.5), providing a quantitative measure of the energy change associated with etching of one layer of Ni atoms from the surface.

$$\Delta G_{\mathrm{rm,layer}} = \frac{1}{x}\{[G_{\mathrm{Ni_s|}} + (\frac{x}{y} - 1)G_{\mathrm{Ni_s|Ni_x}} + xG_{\mathrm{[Ni(L)_2]}} + xG_{\mathrm{H_2O}}] - [\frac{x}{y}G_{\mathrm{Ni_s|Ni_xO_y}} + 2xG_{\mathrm{HL}}]\} \quad (3.5)$$

In the subscripts, following the convention set forth in equation (3.1) - equation (3.2), the species on the left and right of the symbol | refer to the unactivated atomic layers and the activated layers, respectively. The meanings of "$\mathrm{Ni_s}$" and "$\mathrm{Ni_x}$" are already explained above. "$\mathrm{Ni_xO_y}$" refers to the slab activated by y oxygen modifier atoms. The molecular terms are self-explanatory and the same values are used as in the bulk model. The terms in equation (3.5) can be regrouped to yield equation (3.6), in which the three terms can be assigned physical meanings: The first term on the left ($\frac{1}{x}[G_{\mathrm{Ni_s|}} - G_{\mathrm{Ni_s|Ni_x}}]$) differs only in the number of layers included in the slab model, and can be replaced by $-\mu_{\mathrm{Ni,bulk}}$ if the surface formation energy (equation (3.10)) is converged with respect to the number of layers. The second term ($\frac{1}{y}[G_{\mathrm{Ni_s|Ni_xO_y}} - G_{\mathrm{Ni_s|Ni_x}}]$) corresponds to the energy of the modified surface per modifier atom, and is similar to the adsorption energy of the modifier (equation (A.2)) to a constant term. The third term ($[G_{\mathrm{[Ni(L)_2]}} + G_{\mathrm{H_2O}} - 2G_{\mathrm{HL}}]$) describes the bond breaking and formation of the etchant, hydride, and metal complex. For a given substrate/modifier/etchant combination, the first and last terms are constants. The information about activation (ion species, coverage, etc) is expressed through the second term only, thus decoupled from that of the etchant. It is shown that the number of atoms (x) is canceled out, and there is no need to specify the number of atoms per layer.

$$\Delta G_{\mathrm{rm,layer}} = \frac{1}{x}[G_{\mathrm{Ni_s}} - G_{\mathrm{Ni_s|Ni_x}}] - \frac{1}{y}[G_{\mathrm{Ni_s|Ni_xO_y}} - G_{\mathrm{Ni_s|Ni_x}}] + [G_{\mathrm{[Ni(L)_2]}} + G_{\mathrm{H_2O}} - 2G_{\mathrm{HL}}] \quad (3.6)$$

While the layer models are derived for the case of deficient modifier coverage (less than the bulk model reaction stoichiometry), there is no limit on their applicability since the number of metal atoms per layer does not appear in the energy expression. In fact, the layer model reduces to the bulk model if the same assumptions are reimposed. In equation (3.6), the molecular term is the same as that in the bulk model. In the case of bulk NiO, $x = y$,

therefore, the left two terms can be re-written as:

$$\frac{1}{x}[G_{\mathrm{Ni_s}} - G_{\mathrm{Ni_s|Ni_x}}] - \frac{1}{y}[G_{\mathrm{Ni_s|Ni_xO_y}} - G_{\mathrm{Ni_s|Ni_x}}] = \frac{1}{x}[G_{\mathrm{Ni_s}} - G_{\mathrm{Ni_s|Ni_xO_x}}]$$

Since the activation layer is assumed to take the form of the bulk oxide,

$$G_{\mathrm{Ni_s|Ni_xO_x}} \approx G_{\mathrm{Ni_s|(NiO)}_x}$$

following the bulk model, a thick oxide layer is assumed so that the interface and surface formation energies are negligible compared with the energy to etch,

$$\frac{1}{x}[G_{\mathrm{Ni_s}} - G_{\mathrm{Ni_s|(NiO)}_x}] \approx -G_{\mathrm{NiO,bulk}}$$

and the bulk model energy expression is immediately recovered by adding back the molecular terms. However, it is important to note that the layer model on the pristine surface does not fully match the bulk model at low coverages, as it is restricted to hydride formation, while the pristine surface bulk model forms H_2 molecules. This observation may suggest that the layer model is more applicable at significant coverages. However, this aspect was not pursued.

The configurations of the adsorbates are obtained by using a greedy-search algorithm that proceeds iteratively. At each iteration, one adsorbate is added to each available adsorption site, forming a pool of candidate structures for the next round. The adsorption sites themselves are determined by symmetry on the primitive pristine surface unit cell: every unique atom defines a top site, every unique pair of atoms defines a bridge site, and so on. The supercell is not varied. Figure A.2 gives the supercell used and the adsorption sites. Each candidate is then fully relaxed, and the most stable candidate is selected to enter the next round. Figures A.7 to A.10 gives the structures of the selected candidates at every round. The configuration search determines a unique configuration for every termination at discrete coverages. The method adopted here is a static one, meaning that effect of finite temperature is included in the entropy contribution to the free energy, obtained via the calculation of the surface phonon frequencies. A full computational treatment of the temperature effect on activation, requiring dynamic calculations (*e.g.* as molecular dynamics simulation of ion impact), is presented later in chapter 6.

While the energy expression determines the reaction energy normalized to one Ni atom removed (ΔG_{rm}) for a given termination, the results for different terminations span a large range of values at any coverage. As a descriptor for process feasibility, a single-valued prediction is desirable. To minimize the experimental input required, a weighted average of the per-termination etching energies is calculated, with weights given by the equilibrium crystal shape construction (ECS) (equation (3.7)). The Wulff construction is applied to the activated surfaces so that more stable structures with a low surface formation energy are preferred over less stable ones and are assigned a larger weight (a^i).

$$\Delta \bar{G}_{\mathrm{rm}} = \sum_{i \in \{(100),(110),\ldots\}} a^i \Delta G^i_{\mathrm{rm}} \tag{3.7}$$

The determination of a^i requires the surface energies, which, in turn, depend on the modifier coverage, the chemical potential of the modifier, and the substrate. Thermodynamically, the substrate is connected to a large reservoir of pristine atoms, hence (e.g., for Ni):

$$\mu_{\mathrm{Ni}} = G_{f,\mathrm{Ni}} = \text{Gibbs free energy of formation of fcc Ni bulk, per Ni atom} \tag{3.8}$$

The chemical potential of the modifier is difficult to determine confidently due to the various excited species present. Plasma diagnostics or simulation of ion energy distributions (IED) is not attempted. Instead, the assumption of surface-plasma equilibrium allows us to indirectly determine the chemical potential through the simulated modifier coverage (area density). At the coverage observed (the coverage that self-limited adsorption stops at), no termination should have a positive adsorption energy. The chemical potential of the modifier is thus chosen to be the minimum required to set the adsorption free energy of all terminations greater than or equal to 0, leading to a modifier chemical potential as a function of coverage (equation (3.9)).

$$\mu_{\mathrm{O}}(\theta_{\mathrm{O}}) = \min_{i \in \{(100),(110),\ldots\}} \frac{G^i_{\mathrm{Ni_s|Ni_xO_y}}(\theta_{\mathrm{O}}) - G^i_{\mathrm{Ni_s|Ni_x}}}{y} \tag{3.9}$$

In effect, substituted the chemical potential, the usual control parameter for reaction thermodynamics, has been substituted with the experimentally observable parameter of modifier area density (coverage). The surface energies can be calculated by equation (3.10), where γ^i_{Ni}

represents the surface energy of the pristine substrate (needed to correct for the bottom pristine surface exposed in asymmetric slab structures), and s is the number of metal atoms in the pristine substrate. The resulting surface energies are shown in figure A.5. The set of surface energies for a given coverage now allows the Wulff construction to be performed, which minimizes the total surface energy of exposed terminations on crystal grains constrained by a fixed volume. The fractions (a^i) resulting from the minimization (see figure A.6) are used as weights for averaging the per-termination removal step energies.

The rationale behind using the Wulff construction is that terminations more favorable to etching are also less stable and hence may be present in less significant fractions in the ensemble. Admittedly, the ECS method is not entirely consistent with our process, since it minimizes the total surface energy subject to a fixed grain volume, whereas the surface faceting behavior is limited by a fixed cross-sectional area. Hence, no claim is made for the ECS-based model to accurately describe the microfaceting that might occur in response to adsorbate coverage changes. Solving the cross-sectional area constrained minimization problem is outside the scope of the present work, and the ECS method is adopted as a physics-inspired way to systematically assign the weights. Nor is it claimed that microfaceting reconstruction would happen at the target process temperature. The inspiration came from the transient heat spikes present locally near the ion impact region during plasma oxidation, which can temporarily bring the local temperature close to the melting point. Nevertheless, the ECS is an *ad hoc* addition to the model to allow for calculation of single-valued etching energies that can be more readily compared between different proposed ALE processes. Since the per-termination removal step energies are independent of this assumption, when more reliable information on the surface is available (*e.g.*, via the crystal orientation distribution function from X-ray diffraction), the weights can be substituted in.

$$\gamma^i_{\mathrm{Ni_s|Ni_xO_y}} = [G^i_{\mathrm{Ni_s|Ni_xO_y}} - (x+s)\mu_{\mathrm{Ni}} - y\mu_{\mathrm{O}}]/A_0 - \gamma^i_{\mathrm{Ni}} \tag{3.10}$$

3.3 Computational Details

All electronic energy calculations are performed with density functional theory as implemented in the Vienna ab initio Simulation Package (VASP).[27,34,35] The electron-ion interactions are treated using the projector augmented wave (PAW) method[48] and the valence one-electron functions are developed on a basis set of plane waves. The Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional[36] is used throughout. The van der Waals correction to the PBE functional was tested, for example, with the DFT-dDsC method,[92,93] but not applied due to the overestimated bulk formation energies (see appendix A.2). The bulk crystal parameters are obtained by starting from their experimental values[1,25,61,115] and performing a two-step direct volume relaxation.

The configuration search is done in two steps. A reasonably crude but fast computational setup is used to relax all candidate structures and a slow but accurate setup is used for single-point calculations of the most stable structures. The two setups differ in Monkhorst-Pack k-space sampling density with the ($3 \times 3 \times 1$ versus $5 \times 5 \times 1$ mesh) and plane wave cutoff (300 eV versus 400 eV). The relaxed structure and the relative stability of each adsorption site do not change with respect to these parameters. All other parameters are identical. Energies are converged to 1×10^{-6} eV. Forces are converged to 0.02 eV/Å. For the molecules, the plane wave cutoff is maintained at 400 eV.

For the Gaussian[113] calculations on the gas phase molecules, the PBE functional and the triple-zeta split valence basis set of Ahlrichs et. al.(def-TZVP)[24] were used. Geometric structures were converged with forces converged to 1.5×10^{-5} Hartree/Bohr, and displacement were converged to 6×10^{-5} Bohr on an "Ultrafine" grid. These calculations were only used to confirm the relative stabilities and the resulting energies were not used in the reported data.

Free energies reported include translational, rotational, and vibrational contributions to the entropy, whenever appropriate. The phonons of the bulk systems are calculated with density functional perturbation theory (DFPT).[16] The surface vibrational frequencies are calculated for the Γ point only with the bottom two layers frozen. The many degrees of

freedom prevented the practical use of a more accurate setup. The vibrational analysis for molecules are also performed with finite difference, with a step size of 0.015 Å. Soft vibrational modes (resulting from the internal rotations of the Ni(II) complexes) with frequencies smaller than 100 cm^{-1} are shifted to the threshold value. Performing this shift resulted in a change in removal step energy on the order of 0.01 eV to 0.03 eV, depending on the etching chemistry (HFA vs. HFAmd) and the coverage. The free energies are evaluated at 80°C throughout the paper. The pressures are estimated according to the fluxes in the experiment (see equation (A.1)).

The Wulffpack package is used to obtain the evolution of the surface termination distribution[158] with respect to the modifier coverage θ. Since the adsorbates are added to substrates with finite area, the available coverages are discrete. The discrete dataset is interpolated linearly to yield a continuous $\gamma - \theta$ curve, which is then used to perform the Wulff construction. At high coverages, terminations that saturate below a given coverage are not considered and are removed from the Wulff construction calculations (i.e., no extrapolation is performed).

3.4 Results and Discussion

3.4.1 Formate and formamidinate complexes of Ni and Cu

The energetics of the nickel complexes simulated are shown in figure 3.2. For the monomer case, three structures (monodentate, bidentate planar, bidentate tetrahedral) are searched and the bidentate planar structure is found to be the most stable. Our results contradict an earlier study[74,81] that suggested the tetrahedral structure being more stable. We confirmed the trend by performing atomic basis set calculations in Gaussian[113] (see the computational details section). Dimers, trimers, and tetramers of various configurations are built and compared. Three tetramers are built (figure 3.1) and only the most stable structures' energies are plotted in figure 3.2. For the dimers, trimers, and tetramer #3, the chelation angles are the same as the tetrahedral monomer. The bonds from the metal center to the ligand atoms

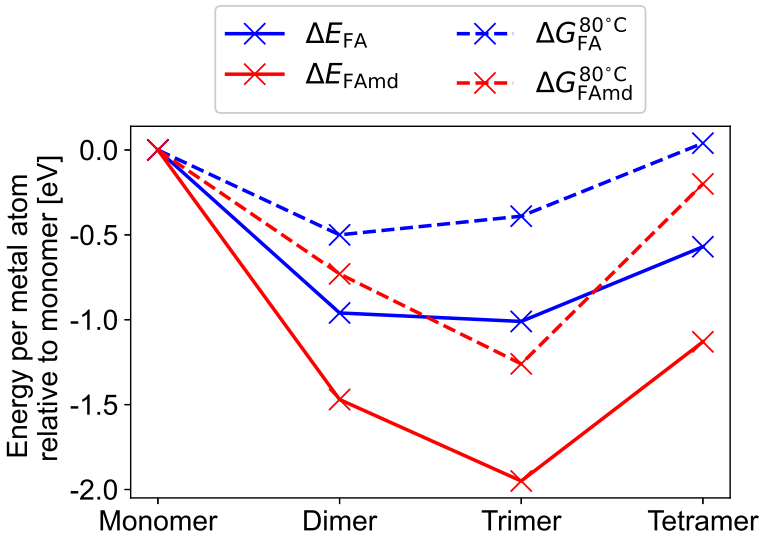


Figure 3.2: Electronic potential energies (ΔE) and Gibbs free energies ($\Delta G^{80\,^{\circ}\mathrm{C}}$) of nickel formate ($Ni(FA)_2$) and formamidinate ($Ni(FAmd)_2$) oligomers (olig.) complexes relative to the monomer (mono.) complex. Energy differences are normalized by nuclearity (nucl.) for comparison on a per metal atom basis.

are progressively elongated with respect to the nuclearity to accommodate more ligands. The formate and formamidinate complexes follow the same pattern, although the metal-nitrogen bonds in amidinate complexes are slightly longer than the corresponding metal-oxygen bonds in formate complexes, presumably due to steric repulsion among the amine groups. By the Gibbs free energy, the most stable formate complex is the dimer, and the most stable formamidinate complex is the trimer. Since thermodynamics cannot preclude kinetic barriers that may prevent the formation of these large complexes, the main results are recalculated with the monomer complex as the product in Figures figure A.11 and figure A.12.

3.4.2 The bulk model

The results from the bulk model on the pristine ($\Delta G_{\mathrm{rm,pris}}$) and activated ($\Delta G_{\mathrm{rm,act}}$) substrates are shown in table 3.1 and table 3.2, respectively.

Both Ni and Cu in their pristine forms are found to resist etching by HFA and HFAmd

due to the large positive values of $\Delta G_{\mathrm{rm,pris}}$, making self-limited reactions possible. The results on bare metal substrates involving HFAmd are expected since substituted derivatives of the formamidinate complex are used to deposit Cu and Ni in atomic layer deposition (ALD).[64] The reducing agent used is H_2, hence the ALD reaction is the exact reversal of the ALE reaction. Overall, Ni has a higher (more unfavorable) removal step energy than Cu for the pristine surface.

Table 3.1: Gibbs free energy of the removal step reactions (reaction {3.1a} - reaction {3.1b}) on pristine substrates from the bulk model ($\Delta G^{80\,^\circ\mathrm{C}}_{\mathrm{rm,pris}}$, equation (3.1)). The energy values are normalized to one metal atom etched.

Substrate	Etchant	$\Delta G^{80\,^\circ\mathrm{C}}_{\mathrm{rm,pris}}$[eV]
Ni	HFA	1.00
Ni	HFAmd	0.76
Cu	HFA	0.50
Cu	HFAmd	0.44

Table 3.2: Gibbs free energy of the removal step on activated substrates (reaction {3.2a} - reaction {3.3b}) from the bulk model ($\Delta G^{80\,^\circ\mathrm{C}}_{\mathrm{rm,act}}$, equation (3.2)). The energy values are normalized to one metal atom etched.

Substrate	Modifier	Etchant	$\Delta G^{80\,^\circ\mathrm{C}}_{\mathrm{rm,act}}$[eV]
Ni	N	HFA	0.34
Ni	N	HFAmd	0.10
Ni	O	HFA	0.26
Ni	O	HFAmd	0.02
Cu	N	HFA	−0.15
Cu	N	HFAmd	−0.22
Cu	O	HFA	−0.06
Cu	O	HFAmd	−0.12

The results on activated surfaces show a sharp contrast between Cu and Ni. Throughout

the text, the units for reaction energies are eV per metal atom etched, unless otherwise stated. All four modifier and etchant combinations give a slightly favorable (negative) removal step energy on Cu, whereas on Ni the same four combinations are unfavorable (positive). The calculated $\Delta G_{\mathrm{rm,act}}$ is 0.10 eV for the HFAmd etching of Ni_3N, which is consistent with reports of Ni_3N deposition using amidinate complexes of nickel (which can be subsequently reduced by H_2 to give Ni films), although the ligands in the report are heavily substituted.[86] We highlight the fact that etching O-activated Ni with formic acid is predicted to be unfavorable by 0.26 eV, inconsistent with the experimentally observed etching.[162]

The bulk model results suggest that formamidine etching is preferable compared to formic acid by 0.24 eV and 0.07 eV, respectively on Ni and Cu. The differences between oxygen and nitrogen activation (0.08 eV favoring oxygen on Ni, 0.09 eV favoring nitrogen on Cu) are not significant.

To better understand the trend, energy decomposition is performed. The decomposed reactions (reaction {3.5a} - reaction {3.6b}) on the pristine surfaces separate the influence of the metal substrate from that of the etchant. In table 3.3, it can be seen that it is more costly to atomize Ni than Cu, yet Ni binds the complexes more strongly. The two effects cancel somewhat, and the overall result is that Ni is less favorable to etch. A similar conclusion can be drawn from the energy decomposition of the activated substrates (reaction {3.7a} - reaction {3.10b}) where the activation with nitrogen or oxygen is separated from the binding of the metal to the complexes and the formation of hydrides. Table 3.4 shows the results. First, note that the oxides are thermally stable with respect to O_2, while the nitrides are not stable with respect to N_2. Second, activation with both N and O are more favorable on Ni (*e.g.*, by comparing Ni/N/FA and Cu/N/FA, Rxn. a) relative to Cu), effectively stabilizing the Ni metal. Since the relative differences on the b reactions for each set are the same as that of the pristine metal (which already indicates that pristine Cu is more favorable to etch), this further stabilization makes the nickel nitride less favorable to etch compared to Cu_3N. The same can be said of the oxides, hence giving the global qualitative difference between Ni and Cu, Ni being harder to etch.

The role of oxidation states can be seen clearly in the data. The fact that reaction {3.8a}

on NiO is 0.93 eV less favorable than reaction {3.10a} on Cu_2O can be partially attributed to the difference in oxidation state change. This is evident since the difference for the nitrides, between reaction {3.7a} on Ni_3N and reaction {3.9a} on Cu_3N, is only 0.15 eV, much smaller compared with that on the oxides.

On the other hand, note that the decomposed reactions cannot be used to compare N and O activation (e.g. by comparing reaction {3.7a} to reaction {3.8a}), as the reference points to N_2 and O_2 are arbitrary and they bear no significance for the actual plasma activation process. The decomposed reactions obviously do not correspond to the actual reaction mechanisms, and their sole purpose here is to separate the influences of the substrate, the modifier, and the ligand on a thermodynamic level.

$$\mathrm{Ni(bulk)} \longrightarrow \mathrm{Ni(atom)} \tag{3.5a}$$

$$\mathrm{Ni(atom)} + 2\,\mathrm{HL} \xrightarrow{\mathrm{NiL_2}} + \mathrm{H_2} \tag{3.5b}$$

$$\mathrm{Cu(bulk)} \longrightarrow \mathrm{Cu(atom)} \tag{3.6a}$$

$$\mathrm{Cu(atom)} + \mathrm{HL} \longrightarrow \frac{1}{2}\,[\mathrm{Cu_2L_2}] + \frac{1}{2}\,\mathrm{H_2} \tag{3.6b}$$

$$\frac{2}{3}\,\mathrm{Ni_3N} \longrightarrow 2\,\mathrm{Ni(bulk)} + \frac{1}{3}\,\mathrm{N_2} \tag{3.7a}$$

$$2\,\mathrm{Ni(bulk)} + \frac{1}{3}\,\mathrm{N_2} + 2\,\mathrm{HL} \xrightarrow{\mathrm{NiL_2}} + \frac{2}{3}\,\mathrm{NH_3} + \mathrm{Ni(bulk)} \tag{3.7b}$$

$$\mathrm{NiO} \longrightarrow \mathrm{Ni(bulk)} + \frac{1}{2}\,\mathrm{O_2} \tag{3.8a}$$

$$\mathrm{Ni(bulk)} + 2\,\mathrm{HL} + \frac{1}{2}\,\mathrm{O_2} \xrightarrow{\mathrm{NiL_2}} + \mathrm{H_2O} \tag{3.8b}$$

$$\frac{1}{3}\,Cu_3N \longrightarrow Cu(bulk) + \frac{1}{6}\,N_2 \quad \{3.9a\}$$

$$Cu(bulk) + HL + \frac{1}{6}\,N_2 \longrightarrow \frac{1}{2}\,[Cu_2L_2] + \frac{1}{3}\,NH_3 \quad \{3.9b\}$$

$$\frac{1}{2}\,Cu_2O \longrightarrow Cu(bulk) + \frac{1}{4}\,O_2 \quad \{3.10a\}$$

$$Cu(bulk) + HL + \frac{1}{4}\,O_2 \longrightarrow \frac{1}{2}\,[Cu_2L_2] + \frac{1}{2}\,H_2O \quad \{3.10b\}$$

Table 3.3: Energy decomposition of bulk model on pristine surfaces. Steps a, b correspond to the labels in reaction {3.5a} - reaction {3.6b}, and refer to bulk metal evaporating and complex/H_2 formation, respectively. The energy values are normalized to one metal atom etched.

Substrate	Etchant	$\Delta G_a^{80\,^\circ C}$[eV]	$\Delta G_b^{80\,^\circ C}$[eV]
Ni	HFA	4.19	−3.23
Ni	HFAmd	4.19	−3.47
Cu	HFA	3.05	−2.55
Cu	HFAmd	3.05	−2.62

3.4.3 The layer model - per termination

The fact that HFA is able to etch nickel films activated with oxygen plasma suggests that the substrate resulting from the plasma activation could be quite different from the bulk nickel oxide. The predicted unfavorable Gibbs free energy for the removal step of 0.26 eV is significant enough to conclude that the bulk model is inadequate to offer qualitatively correct predictions for the plasma-activated surface. While the bulk model lacks the flexibility to account for the variety of surface structures, as it is limited to the few known bulk lattices, its simplicity makes it useful as a reference and sanity check for the more sophisticated layer model.

Table 3.4: Energy decomposition of bulk model on activated surfaces. Step a, b correspond to the labels in reaction {3.7a} - reaction {3.10b}, and refer to reverse of oxidation/nitridation (modifier desorption) and complex/hydride formation, respectively. The energy values are normalized to one metal atom etched.

Substrate	Modifier	Etchant	$\Delta G_{\mathrm{a}}^{80\,^{\circ}\mathrm{C}}$[eV]	$\Delta G_{\mathrm{b}}^{80\,^{\circ}\mathrm{C}}$[eV]
Ni	N	HFA	−0.37	0.71
Ni	N	HFAmd	−0.37	0.47
Ni	O	HFA	1.45	−1.19
Ni	O	HFAmd	1.45	−1.43
Cu	N	HFA	−0.52	0.37
Cu	N	HFAmd	−0.52	0.31
Cu	O	HFA	0.52	−0.58
Cu	O	HFAmd	0.52	−0.64

The layer model energy expression given in equation (3.6) relies on the adsorption energy as the central coverage-dependent term. As the adsorption energy is the more established metric to quantify adsorbate-surface interaction, the sampling of the surface structures is discussed in terms of it. The adsorption energies of the structures picked by the configuration search are shown in figure 3.3. The reference points are arbitrarily chosen to be gas phase O_2 and N_2 for O and N adsorption, respectively. This choice does not affect the rest of the model. The general trend in all four combinations is that the adsorption energies tend to increase (become less favorable) with increasing coverage as the more favorable sites are preferentially occupied by our algorithm. The lateral repulsion among adsorbates at high coverages also makes the high-coverage configurations less favorable. However, on the Cu/N system, the trend is not completely followed. In figure 3.3, around $0.06\,\text{Å}^{-2}$, all adsorption energies on all terminations except (100) decreased to form a "dip". This is explained by referring back to the structures shown in figure A.8. On the Cu (111) termination, the most significant "dip" occurred between structures with 2 and 3 nitrogen atom adsorbates. On the 3-adsorbate structure, the surface in the vicinity of one of the adsorbed nitrogen atoms

relaxed into a square pseudo-(100) structure, as previously reported both experimentally[47] and computationally.[73] The 4-adsorbate structure has two adsorbates forming the pseudo-(100) structure. Eventually, the increasing trend resumed due to inter-adsorbate repulsion as the adsorbates become closer. It also explains why the "dip" is not observed on the (100) termination. The (211) and (221) terminations follow a similar trend due to the large (111) terraces present.

The adsorption energy of nitrogen is $\sim 1.5\,\mathrm{eV}$ less favorable on Cu than on Ni, while oxygen adsorption is $\sim 0.5\,\mathrm{eV}$ less favorable. In particular, nitrogen adsorption is unfavorable with respect to N_2 at all coverages for Cu surfaces and at coverages above $\sim 0.10\,\mathrm{\AA}^{-2}$ on Ni surfaces, suggesting that these surface nitrides on Cu and Ni are not thermodynamically stable. Without significant energy barriers, the surface nitride may spontaneously decompose to form N_2 and the pristine surface. The practical implication is that if nitrogen is used as the modifier, the activated surface must present a significant kinetic barrier to prevent desorption once the nitrogen plasma is purged. The O adsorption on both substrates is favorable with respect to O_2 for all terminations and coverages considered.

The adsorption energies of oxygen on Cu and Ni in some experimentally known low-coverage configurations are indicated in figure 3.3. The complete list of structures investigated is shown in table A.3. The comparison reveals that the adsorption energy trajectories discovered by the configuration search lie very close to the experimentally observed surface structures, as can be seen for the Ni(100), Ni(111), and Cu(100) terminations. The surface supercell used is briefly investigated for the (100), (111), and (211) surfaces by considering a few supercells with symmetries different from the ones used for the configuration search. The results (see figure A.3) indicate that the errors introduced by supercell selection are small. Furthermore, in the event that the configuration search misses some more stable structures at a given coverage, this error does not propagate significantly through the Wulff construction to the final, averaged removal step energy. For structures with extended reconstruction (e.g., missing row (MR) and pairing-row (PR) reconstructions on the (110) termination), the adsorption energies are much more favorable than found here. Since these structures require the addition or removal of metal atoms, our algorithm cannot recover them. De-

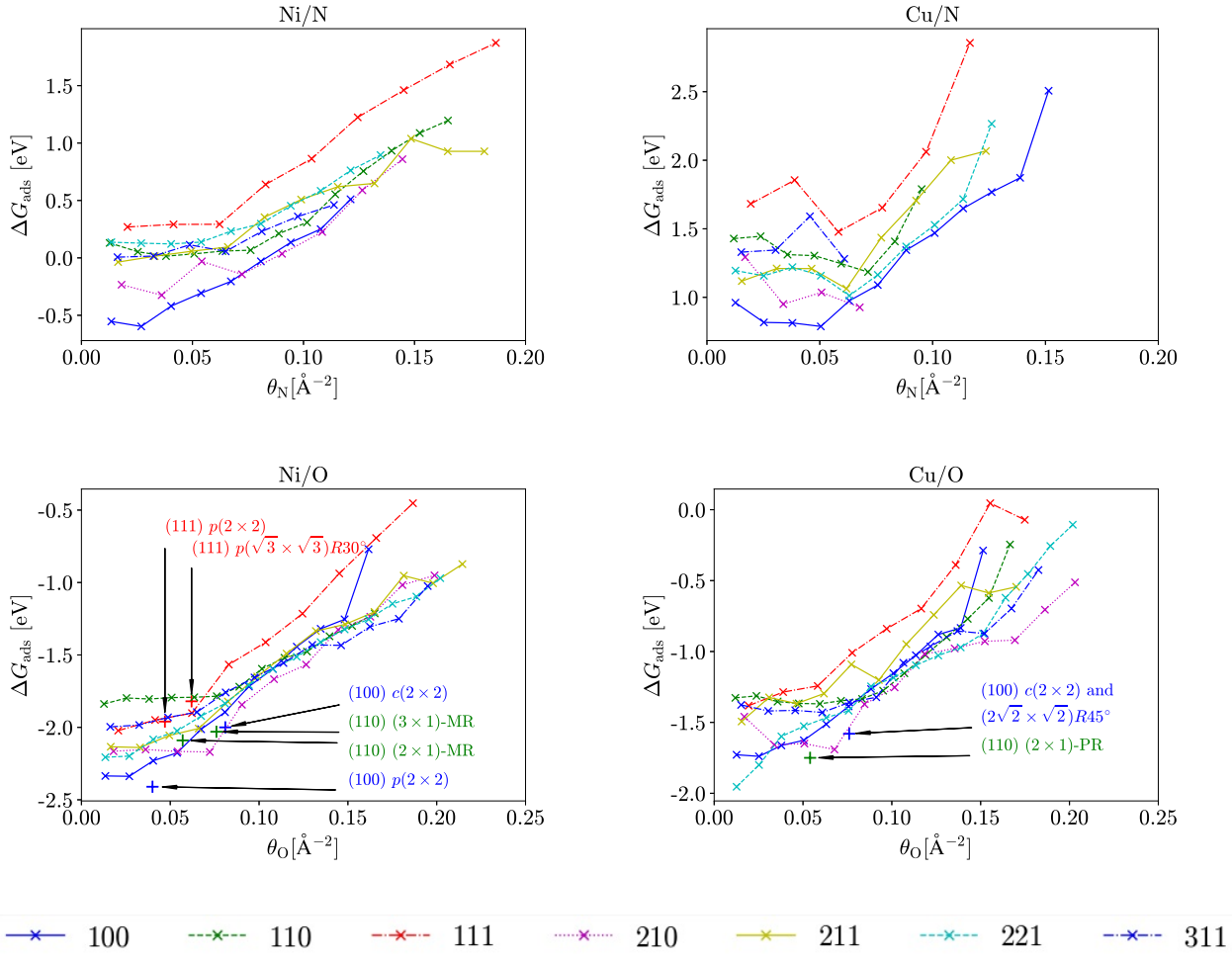


Figure 3.3: Adsorption energies of different terminations of the Ni/N, Cu/N, Ni/O, Cu/O systems as a function of coverage. The labels "PR" and "MR" represent "pairing row" and "missing row" reconstructions of the (110) surfaces, respectively.

pending on the size of the crystalline domains on the amorphous or polycrystalline surface, these reconstructed domains may or may not be important. The comparison indicates that our configurations lie close to the global minimum when extended reconstructions are not considered.

The removal step energies for each surface termination are shown in figure 3.4. The formic acid and formamidine results are shown on the left and on the right y-scales, respectively. The results from the bulk model are indicated with horizontal lines. Unlike the adsorption energies, here the absolute scales are meaningful.

For the Ni surfaces, the layer model produces removal step energies similar to those of the pristine bulk model at low coverages. As the coverages increase, all terminations become more favorable to etch than the pristine bulk. The curves of different terminations decrease with roughly the same slope, but some terminations saturates (terminate) at lower coverages. At the highest coverages explored (the "saturation coverage"), ΔG_{rm} is slightly below 0.0 eV for all terminations except (110) which saturated just above 0.0 eV. In particular, on Ni (111) surfaces, at the coverage of 0.18 $\mathrm{\AA}^{-2}$, etching is the most favorable (-0.50 eV and -0.75 eV for Ni/N and Ni/O with formic acid, respectively) among all terminations investigated.

For the Cu systems, all terminations lie below the line of the bulk metal (figure 3.4 b & d, solid black line). The curves do not have an easily recognizable slope due to reconstructions discussed earlier. The terminations for the Cu/N system saturate early compared with those in the other three systems, suggesting nitrogen-based modification may not persist long enough for thermal activation, as noted earlier. As with the Ni surface, most of the curves crossed the 0 eV line, resulting in favorable etching at high enough coverage.

3.4.4 The layer model - ECS averaged

While a lot of information is presented in the per-termination ΔG_{rm}, it is difficult to relate it to the experimental data. At a given coverage, the difference between the most favorable and least favorable terminations is on the order of 0.5 eV for every system. To proceed, one must average the removal step energies on the different occurring terminations. The present

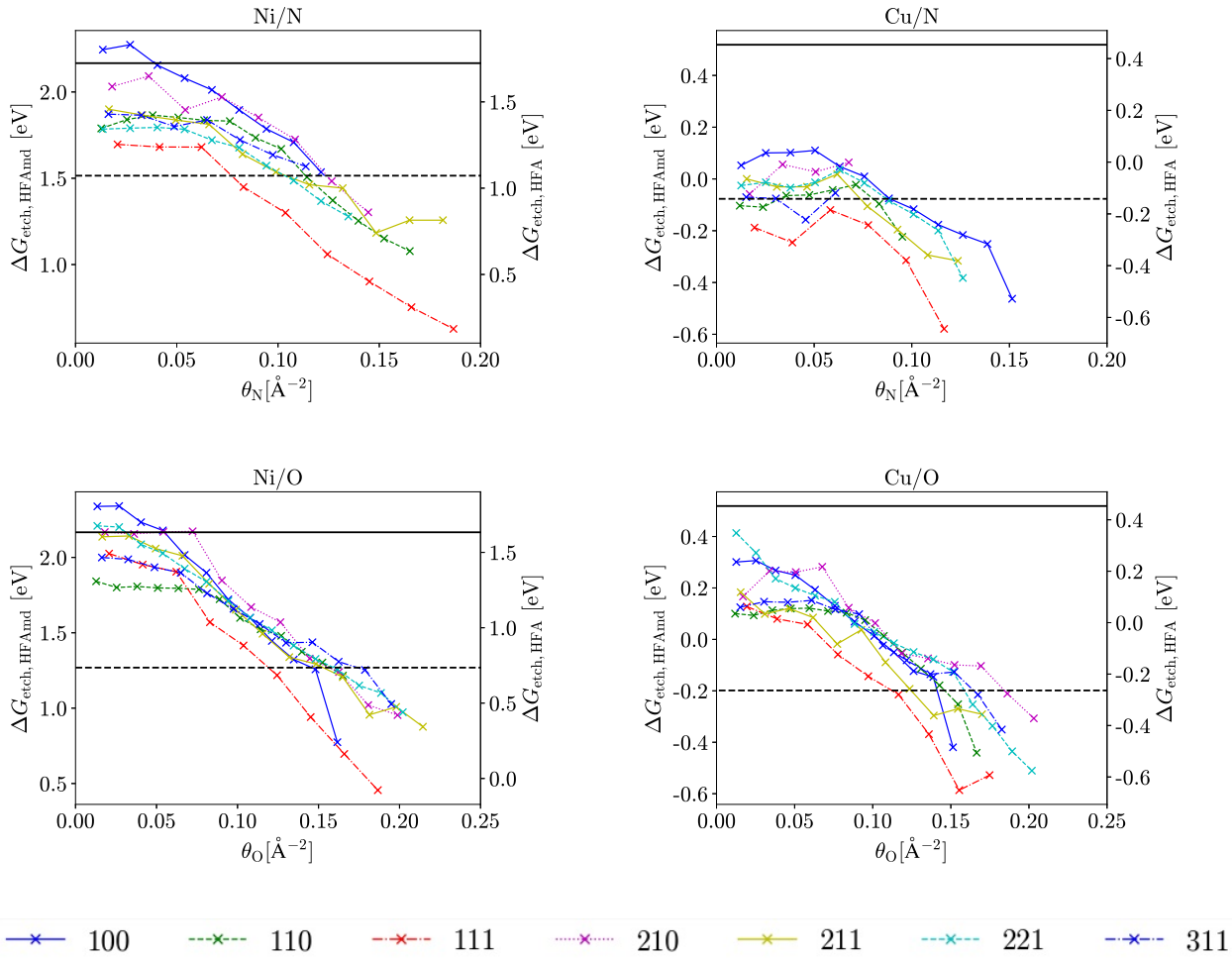


Figure 3.4: Removal step energies by terminations on the Ni/N, Cu/N, Ni/O and Cu/O systems as functions of coverage, evaluated at 80 °C. Two axis are shown. The y-axis on the left corresponds to formamidine chemistry. The y-axis on the right corresponds to formic acid chemistry. The bulk model results on the pristine and activated substrates are respectively marked with colored and black horizontals lines for comparison. The values are in eV per metal atom etched.

work relies on the equilibrium crystal shape with the inverse definition of modifier chemical potentials, as described in the methods section. The chemical potential determined for four pairs of substrates and modifiers are shown in figure A.4. The resulting surface energies and fractions of surfaces are shown in figure A.5 and figure A.6, respectively. The average removal step energies $\Delta\bar{G}_{\mathrm{rm}}$, the final result of the layer model calculated via equation (3.7), are given in figure 3.5. The results from the bulk model (both pristine and activated) are also provided for comparison.

To illustrate the effect of the surface-averaging method, compare figure 3.4(a) to figure 3.5. On the Ni substrate, both O and N modifiers show a significantly more favorable etching on the (111) termination compared with the rest. However, the averaged curves in figure 3.5 do not follow this trend indicated by the (111) termination due to its diminishing surface fractions (see figure A.5 - figure A.6) according to the ECS method. Above $0.08\,\mathrm{Å}^{-2}$ on Ni/N and $0.10\,\mathrm{Å}^{-2}$ on Ni/O, the (111) termination is essentially absent, and the averaged removal step energies are dominated by the higher-indexed terminations. The method therefore prevents biasing the final averaged removal step energy toward any particular termination.

The Ni/N and Ni/O curves follow a smooth, nearly linear decrease with respect to coverage. Both rapidly decrease below the bulk oxide and nitride lines. Over the coverage investigated, the bulk model results lie near the center of the range of removal step energies found. It can be seen that the bulk model predictions lie close to the center of the range spanned by the layer model predictions. The N and O curves give very similar energetics in the middle of the range observed, and both cross the $\Delta\bar{G}_{\mathrm{rm}} = 0\,\mathrm{eV}$ line to become favorable for etching at $\sim 0.14\,\mathrm{Å}^{-2}$. Thus, our layer model predicts that with overlayer adsorption structures and oligomer product structures, etching on Ni can become favorable already. The experimental fact is that O_2 plasma-activated Ni can be favorably etched with formic acid at 80 °C. Our results thus qualitatively agree with the experiments, despite significant difference in activation layer thickness.

On the Cu substrates, the most striking observation is that the two curves of Cu/N and Cu/O are nearly parallel to each other. This is especially surprising when one realizes that

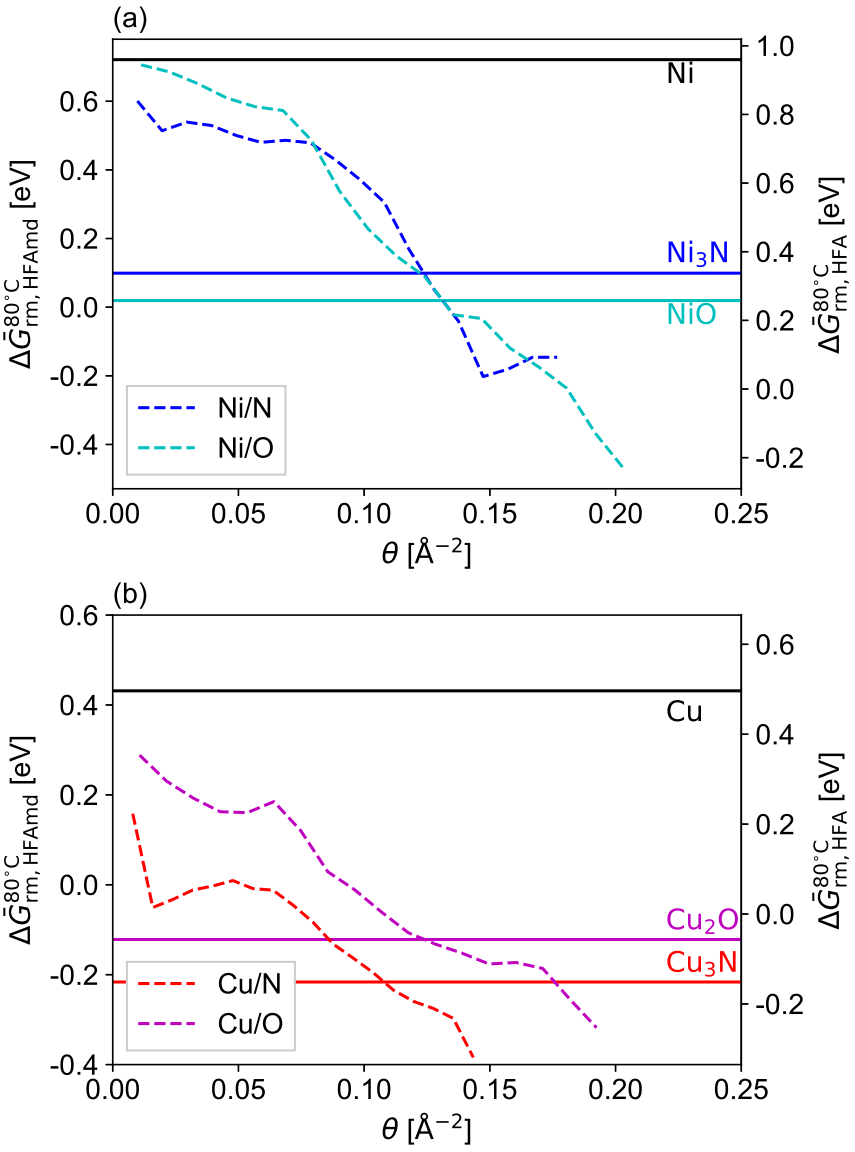


Figure 3.5: Wulff construction-averaged removal step energies as functions of coverage (dashed lines) for (a) nickel and (b) copper systems. The solid lines correspond to the results from the bulk model for the pristine (black) and modified (colored) systems. The values are in eV per metal atom etched.

in figure A.6 the parallel curves are results of contributions from different terminations. The Cu/N curve mostly consists of contributions from (100) and (221) terminations, where as on the Cu/O curve, the (100) is suppressed at low coverages, and the significant contributors are (210) and (211). Our results clearly indicate that the energetics of surface adsorption structures are similar between N and O adsorbates, as was previously demonstrated by a DFT study.[79]

Additionally, the layer model predicts that etching N-activated Cu stays $\sim 0.2\,\text{eV}$ more favorable than O-activated Cu, consistent with the bulk model prediction of $0.09\,\text{eV}$ more favorable etching with N. This can be understood in terms of the difference in bonding: Cu_3N has 1/3 Cu-N bonds per Cu atom, while Cu_2O has 1/2 Cu-O bonds per Cu atom. Since etching 1 Cu atom leads to 1 O-H/N-H bond being formed in both cases, etching Cu/N breaks fewer bonds in total, and a more favorable removal step energy is expected. The same cannot be said of the Ni etching due to its nitride and oxide having different valences.

As shown earlier, the layer model reduces to the bulk model and would give identical removal step energies if additional assumptions of the bulk model are imposed. Our results show that the energetics of etching oxide/nitride on top of the pristine crystal are different from that predicted by the bulk model. Indeed, the integral heats of adsorption at high coverages of oxygen on Ni have been determined to be much lower than the bulk NiO formation energy.[43]

The most useful results obtained are as follows:

Nitrogen activation could lead to working processes similar to oxide (already demonstrated with formic acid) on both substrates, as nitrogen activation gives a more favorable removal step than the oxygen activation.

Formamidine leads to more favorable removal step energies compared with formic acid. This is potentially important as formamidine is a much more volatile etchant than formic acid ($337\,\text{mmHg}$ vs $40\,\text{mmHg}$ at $298\,\text{K}$). Potentially formamidine could yield higher etch rates than formic acid, which is important for industry and for characterizing the etching product.

The trend toward more favorable removal step energies at higher coverages suggests that activation should always aim for higher modifier content.

It is important to note that the configuration search is limited to surface adsorption sites. As shown in chapter 4, following the general trend of more favorable etching at higher coverages, the inclusion of subsurface sites into the configuration search scheme continues this trend and yield more favorable energetics.

This study also reveals that the viability of the ALE process is determined by several factors. As expected, the cohesive energy of the metal plays a significant role. The easier it is to atomize the metal crystal, the more favorable the etching would become. The relative strength of metal-modifier, hydrogen-modifier, and hydrogen-hydrogen bonding, as well as the number of bonds broken and formed, also influence the removal step energy outcome. In particular, the redox potential of the metal plays a role in the strength of the metal-modifier bond, as this interaction involves a change in the metal oxidation state. Moreover, the difference between the hydrogen-modifier bond energy and the hydrogen-hydrogen bond energy affects the selectivity of etching the activated surface compared to the pristine surface: the stronger hydrogen binds to the modifier compared to another hydrogen atom, the stronger the driving force for the etching reaction on the activated surface, and the more likely the removal step is self-limited. The free energy of the metal complex is also of utmost importance. While the most stable structure can be searched in a straightforward manner as done here, there is little guarantee that the most stable molecule is the actual molecule formed. The case in point is the nickel complexes (figure 3.1). While the dimers, trimers, and tetramers are all thermodynamically more stable than the monomer, the formation of these large complexes on the surface may be prohibitively slow. In such cases, the monomer-based thermodynamics could be more relevant to the design of ALE processes.

The factors discussed above can be estimated easily through simple calculations or reference to standard thermochemical data tables. However, influences related to surface effects, such as evolution of surface morphology on the same termination under modifier adsorption and surface heterogeneity effects cannot be determined *a priori* without a detailed study. The influence of surface morphology is clearly indicated when the layer model (assuming a

mononuclear nickel formate complex) predicts etching with formic acid to be unfavorable on Ni (111), but favorable when the subsurface sites become occupied (as shown chapter 4). The heterogeneity among different terminations must be treated carefully for thermodynamic predictions to guide the experiments. The approach taken in the present work relies on the surface energies to summarize the results from various terminations (higher surface energy structures are suppressed). Whenever experimental crystal orientation distribution functions are available, such information can be directly utilized to yield more accurate pictures. Note that the model here is still an 'miller-index-averaged' energy: within each miller index, the model here does not provide information on which etching site is preferrable, unless specific structures that isolate certain sites are provided as input, as demonstrated in chapter 4.

To summarize, the factors favoring metal-plasma ALE processes are: weaker metal-metal bonding (that still allows self-limited etching), stronger modifier-hydrogen bonding, more stable (small) complexes, and less stable activation layer (that does not desorb spontaneously).

As a final note, it is evident from the results of the layer model that "atomic layer etching" is not removing strictly removing atomic layer for certain substrates. To achieve the precise atomic layer removal, the activation has to be limited to the top layer. However, for some substrates, this constraint might make it difficult to find modifiers and etchants that give a favorable removal step energy at these coverages. Alternatively, etching may stop before all the modifiers are consumed, since the low coverage of modifiers left on the surface would lead to unfavorable etching. Thus, the observed etch per cycle can still be one layer or less but the modification penetrates deeper than that, persisting to the next activation step. In both cases, the etching process influences more than one atomic layer. The reader is reminded that the layer model itself was derived with the hypothetical reaction that leads to complete removal of modifier atoms and exposure of the underlying pristine metal. In essence, it is a thermodynamic average of the removal step, where, as the etch proceeds, modifier coverage continuously decreases. It is not possible to pinpoint the coverage or activation thickness at which the etching stops with the present model. Such a prediction would require equation (3.5) to be re-written with the post-removal surface on the product side, which is outside the scope of this work.

3.5 Conclusions

Atomic layer etching (ALE) is a promising technique that has the potential to revolutionize semiconductor manufacturing processes. It relies on time-separated steps of self-limited surface reactions to achieve high selectivity, specificity, and anisotropy, making it suitable for etching a wide range of materials. However, the vast chemical space involved in ALE, including choices of substrate, modifier, etchant, and process parameters, makes it challenging to explore using a trial-and-error approach.

In this work, the thermodynamics of the removal step is proposed as a descriptor for ALE chemistries. A computational framework has been developed to evaluate this descriptor, requiring only minimal input from experiments. Experimental data on the etch product, crystal termination distribution, and post-plasma surface structure can be used to enhance the model's accuracy.

Two models have been proposed and evaluated: the bulk model and the layer model. The bulk model is simpler but lacks the flexibility to address complex surface effects. To overcome this limitation, the layer model relies on surface adsorption calculations to yield a single-valued removal step free energy change. The layer model study indicates that etching becomes more favorable at higher coverages of modifier atoms. Surface oxides and nitrides at high coverages are more favorable to etch than bulk oxides and nitrides. Nitrogen activation leads to more favorable etching than oxygen activation on Cu surfaces, while Ni surfaces show similar removal step energies for both nitrogen and oxygen activation. The layer model demonstrates qualitative agreement with experimental results, achieved by destabilizing the surface through sublayer adsorption structures or stabilizing the products through polynuclear metal complexes.

As simulating molecular impact at the DFT level remains computationally expensive, the surface activation layers were obtained using the greedy-search algorithm. Free energies were extrapolated from 0 K using ab-initio thermodynamics. Once an atomistic model of plasma activation becomes available, the layer model energy expression can be directly applied to such realistic structures.

The modeling efforts can aid ALE process development in several ways. First, it allows the separation of the effects of etchants from the rest of the process, facilitating rapid screening of etchant structures using high-throughput calculations. Second, it highlights the importance of surface morphology and heterogeneity, calling for diverse sampling of surface structures and physically consistent termination-averaging models. Existing concepts in surface science, such as adsorption energy and surface formation energy, provide additional insights into the parameters controlling the calculated removal step energy descriptor.

CHAPTER 4

Site Specificity of the Thermodynamic Model

4.1 Introduction

As discussed in chapter 1, most reported reactive ion etch (RIE) processes on Si use reactive neutrals (e.g. halogens) to weaken the binding between substrate atoms. Fast but inert ions are then introduced solely as a source of directional kinetic energy to volatilize the activated surface species. The same concept is applied to Si atomic layer etching (ALE), with the difference that the doses of ions and neutrals are now time-separated to achieve digital control of the etch thickness. As the ions are inert, the material removal process is essentially a physical sputtering process. The etching selectivity, in this case, relies on the sputtering threshold difference between the activated surface species and the non-activated parts,[38]

$$Y = A\left(\sqrt{E_{\text{ion}}} - \sqrt{E_{\text{threshold}}}\right) \tag{4.1}$$

where Y is the sputtering yield, $E_{\text{threshold}}$ is the sputtering threshold, and E_{ion} is the ion energy. Using reactive species along with ions change the prefactor A and the threshold energy but the energy dependence remains the same, indicating that a similar physical picture remains true.

This approach has been successful for materials like silicon and oxides. However, for more complex materials such as metal oxides, both etching and deposition can occur simultaneously, and the relative rates of these processes become crucial. The simple dependence on ion energy is lost, especially in the case of metals where reactive ions can modify surfaces at energies below the sputtering threshold. This opens up new possibilities for highly selective etching since the chemical reactivity can be tailored to the substrate species more readily

than just relying on the 1-dimensional ion energy.

This is the physical motivation behind using a low-energy plasma-based modification step combined with a thermal removal step. For etching metals, a significant oxidation state barrier must be overcome. Localizing electrons to the surface layer via oxygen (or other oxidative species) plasma exposure is a viable approach. Subsequently, it has been demonstrated that organic chemistries can selectively eliminate the surface metal oxide layer without reacting with the underlying pristine metal material. This selectivity in etching enables precise control over material removal and opens up various practical applications. In such a process, selectivity is defined solely from the oxidation process, the controllable formation of the surface oxide layer is thus critical in realizing the subsequent material removal, including both selectivity and directionality.

In section 4.2 such a process is demonstrated to show the modification effect of low-energy reactive ions. A thin surface layer of Ni (5 nm) is intentionally transformed from metallic Ni to NiO using low-energy oxygen ion exposure. This transformation creates new reaction pathways specifically for removing NiO without affecting Ni substrate. The chemical contrast between surface NiO and bulk Ni results in nearly infinite etching selectivity, similar to etching SiO_2 and Si using HF. This high selectivity is due to the different reactivity of NiO compared to metallic Ni, allowing NiO to be selectively removed while leaving the Ni substrate intact.[128,129] Note that the experimental work is performed by Sang Xia and Jane P. Chang and reproduced here for completeness.

In chapter 3, a thermodynamic model is developed that yields a coverage-dependent, structure sensitive thermodynamic estimate of the energy to etch the activated film. As noted in passing in that chapter, since this model relies on the total energies on a slab structure, it gives the surface averaged etching energy. In this chapter, the model is adapted to study the effect of site-specificity of the nickel oxide-based ALE process, removing the surface-averaging by constructing structure models that contain only well-defined sites. The modeling results, along with the experimental results, suggest that site-dependency needs to be explicitly accounted for. This motivates the subsequent chapters (e.g., chapter 5 and chapter 6), where the plasma oxidation dynamics are explicitly considered.

4.2 Experimental collaboration

In the assessment of controlled surface oxidation, a low-energy plasma oxidation process was used on 30 nm to 40 nm thin films of Ni. The plasma oxidation was performed at a pressure of 30 m Torr, with a source power of 500 W, and no applied substrate bias or heating, for a duration of 2 min. The surface atomic composition and chemical bonding configuration before and after oxidation were analyzed using X-ray photoelectron spectroscopy (XPS) with a monochromatic Al Kα source and a pass energy of 20 eV. The change in film thickness was measured using scanning electron microscopy (SEM) and transmission electron microscopy (TEM).

Prior to the plasma oxidation, the Ni sample was sputter-cleaned using 4 keV Ar ions to remove the native oxide layer and establish a baseline for compositional analysis. The sample was then exposed to oxygen plasma for 2 minutes. Due to ambient exposure, some adventitious carbon was observed, which caused signal attenuation of other elements. The ion energy during the plasma oxidation was about the plasma self-bias potential, which was approximately 10 eV, lower than the reported sputtering threshold of Ni, which is approximately 16 eV.[14] Thus, the main effect of the low-energy ions was oxidation, and no significant sputtering was expected due to the small populations of ions at the high-energy tail of the ion energy distribution function.

The formation of NiO is confirmed from the *ex-situ* XPS analysis where the signal intensity of metallic Ni (852.6 eV) is considerably reduced and that of NiO (854.1 eV) is greatly increased. Increase in signal intensity is also noticed for $Ni(OH)_2$ (855.8 eV) and satellite features (858.1 eV and 859.9 eV) (see figure 4.1). Using the relation $\lambda = 2170/E^2 + 0.72(aE)^{0.66}$ where E is the electron kinetic energy and a is the molecule size in nanometer as derived from the relation $\rho N_A n a^3 = 1 \times 10^{24}\,\mathrm{A_0}$, where ρ is the material density, n is the atoms per "molecule" and A_0 is the molar mass, the mean freepaths of Ni and O were determined to be 2.1 nm and 1.7 nm, respectively. The oxide thickness was then determined to be 6 nm.

$$I^t_{ox}/I^t_{Nt} = I^\infty_{ox}\left[1 - \exp\left(-\frac{t_{ox}}{\lambda^o_{ox}\sin\theta}\right)\right]/I^\infty_{Ni}\exp\left(-\frac{t_{ox}}{\lambda^{Nt}_{ox}\sin\theta}\right) \tag{4.2}$$

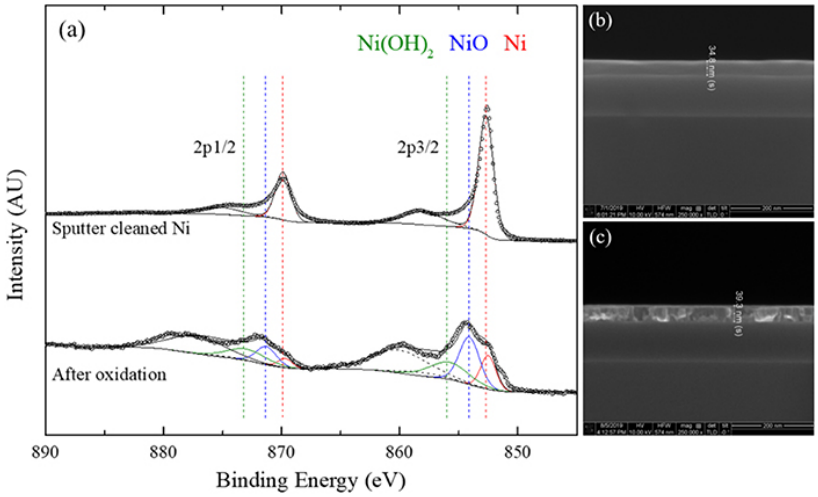


Figure 4.1: (a) *ex-situ* XPS detail scan of Ni 2p and SEM cross-sectional image for Ni blanket sample (b) before and (c) after 2 min of oxygen plasma exposure, 500 W power, and 0 W applied bias.

The calculated NiO thickness obtained from XPS analysis is supported by SEM measurements, which show a thickness increase of approximately 5 nm on the oxidized Ni thin film compared to the non-oxidized surface. It is worth noting that the formation of a native oxide layer of approximately 2 nm takes about 24 h under ambient conditions, which is considerably longer than the time required for the *ex-situ* XPS measurements on the post-oxidation samples.

In Figure 4.2(a), Ni 2p, C 1s, and O 1s spectra are shown with various oxidation times under the same plasma conditions. As the oxidation time increases, more oxides of Ni are formed. The relative concentration between hydroxides and oxides begins to change at longer oxidation times (between 1 min to 4 min). The changes in the amount of metallic, oxide, and hydroxide states of Ni are depicted in Figure 4.2(b).

In Figure 4.3(a), Ni 2p, C 1s, and O 1s spectra are shown with various substrate bias powers applied for 30 s under the same plasma oxidation conditions. With increasing substrate bias power (0 W to 20 W), a nearly monotonic increase in the amount of oxidized nickel is observed. This indicates that the extent of oxidation can be controlled by adjusting the substrate bias power. See chapter 6 for a theoretical model of ion-energy controlled oxide

growth.

The formic acid chemistry in the gas phase is only reactive with oxidized nickel. Therefore, the extent of oxidation achieved during the directional oxidation step translates to the attainable etching rates during the subsequent material removal step, while maintaining very high selectivity. The targeted reaction in the directionally oxidized region enables directional removal, where the reaction viability is determined by the thermodynamics of the reaction, rather than the energy of the ions. *In situ* mass spectrometry analysis was attempted to identify and characterize the gas phase etch products but was not successful due to both the low etch rates and small concentrations of the reaction products. Computational insight is therefore needed to elucidate the reaction pathways and products involved in the ALE process.

Note that here the influence of gas phase products is not explicitly accounted for, but it can be easily re-incorporated since the energy offset between different products is a constant equation (3.6).

4.3 Methods

To provide molecular scale understanding and determine the thermodynamic balance and reaction products during formic acid etching of oxidized nickel, density functional theory (DFT) calculations were performed. The PBE exchange-correlation functional was used with the VASP code.[27,34–36,48] One-electron functions were developed on a basis set of plane-waves, with an energy cutoff at 400 eV. The energies and forces are converged to 1×10^{-6} eV and 1×10^{-2} eV/Å, respectively. The Gibbs free energies of the reaction were evaluated at a specific processing temperature, 80 °C, and at a pressure of 350 Torr.[132]

Three nickel surfaces, (100), (111), and (211), were selected since they are representative of facet and edge sites present on polycrystalline surfaces. To create an oxidized nickel surface, atomic oxygen was chemisorbed on nickel. At the coverages of interest, the most stable phase is surface adsorption (overlayer). The metastable structures likely to be present after oxygen adsorption are probed by placing oxygen in the sublayer interstitial sites (sublayer

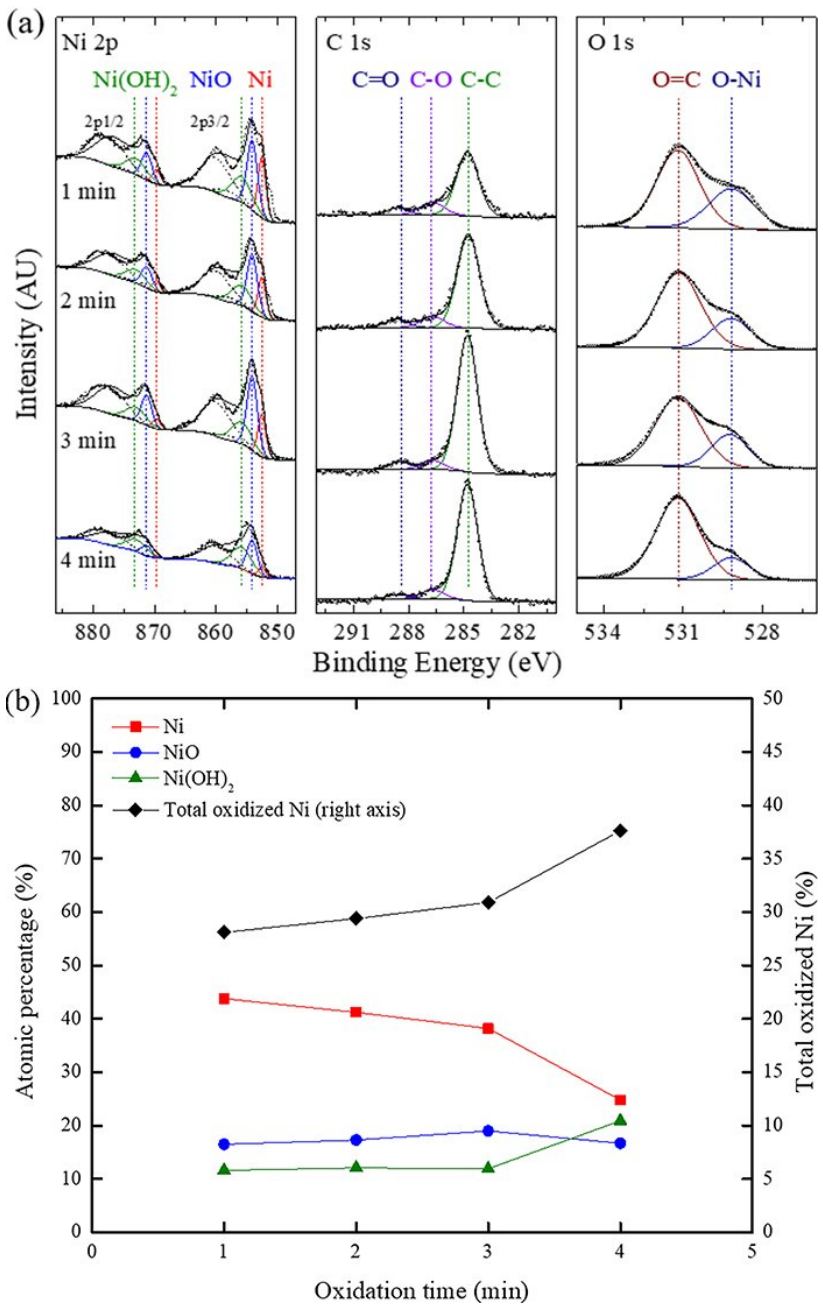


Figure 4.2: (a)*ex-situ* XPS of Ni 2p, C 1s, and O 1s and (b) the relative atomic percentage of different nickel chemical states as a function of oxidation time. All experiments started with 15 s Ar sputtering, followed by in situ oxidation at 500 W power and 0 W applied bias.

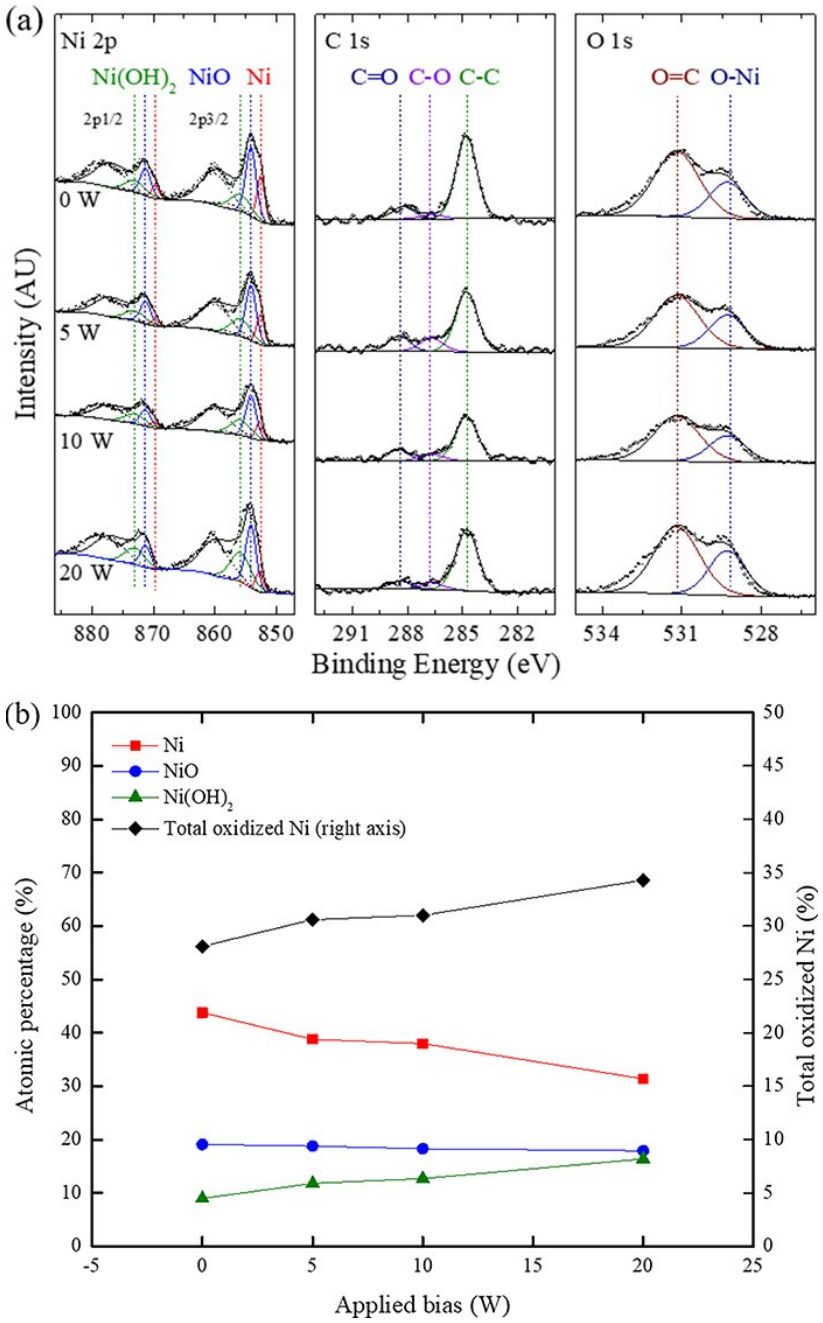


Figure 4.3: (a) XPS of Ni 2p, C 1s, and O 1s, and (b) the relative atomic percentage of different nickel chemical states as a function of substrate bias. All experiments started with 15 s Ar sputtering, followed by in situ oxidation at 500 W power for 30 s.

structures). In a real oxidation process, some overlayer sites are likely to be populated first due to the kinetic barriers of accessing the sublayer sites; therefore, a mixed structure presenting overlayer and sublayer sites is also included. Figure 7 shows the structures for the three surfaces. The sites on the stepped (211) surface deserve some explanation. While the overlayer and mixed structures initially had different site occupation, after relaxation they became very similar. Two types of oxygen atoms are present, namely 4-coordinated (4c) and 3-coordinated (3c). The overlayer structure has one 4c O atom and two 3c O atoms. The mixed structure hastwo 4c O atoms (one in (100) hollow site, one in tetrahedral interstice site) and one 3c atom (in (111) surface site). The sublayer structure has only 3c O atoms.

The PAW basis set employed by the VASP code is cut off at 400 eV. Surface species are calculated with orbital occupancies evaluated using second order Methfessel-Paxton electron smearing at 0.2 eV. Molecular species are calculated with gaussian smearing at 0.02 ev. Geometry optimization were terminated when forces are smaller than 0.02 eV/Å. Electronic steps are converged to 10 eV to 6 eV. Vacuum layers of 16 Å were used to pad the slabs to avoid unphysical slab interactions. Surface dipole corrections were used to obtain correct electrostatic energies. Spin polarization was included for all periodic systems and all molecules except H_2, where spin polarization leads to an incorrect ground state energy.

To study coverage effects, three scenarios are considered here, represented in the chemical reactions below:

$$\text{Pristine Ni} \cdot \text{Ni} + 2\,\text{HCOOH} \rightarrow \cdot \ + [\text{Ni(HCOO)}_2] + \text{H}_2 \qquad \{4.1a\}$$

$$\text{Low O coverage} \cdot \text{NiO}\theta + 2\theta\,\text{HCOOH} \rightarrow \cdot \ + [\text{Ni(HCOO)}_2] + \theta\,\text{H}_2\text{O} \qquad \{4.1b\}$$

$$1\,\text{ML Oadsorption} \cdot \text{NiO} + 2\,\text{HCOOH} \rightarrow \cdot \ + [\text{Ni(HCOO)}_2] + \text{H}_2\text{O} \qquad \{4.1c\}$$

The monolayer (ML) coverage is redefined with respect to the number of nickel atoms exposed to modifiers (O). For p(2×1) cells used in (111) mixed structure, for instance, two nickel atoms are exposed, hence 1ML corresponds to 2 oxygen atoms, one in the overlayer and the other in the sublayer. The θ's of low coverage cases are specified in table 4.1.

Since no experimental characterization of the gas phase nickel di-formate complex $[\text{Ni(HCOO)}_2]$ is available, various structures were considered with DFT, and the most stable one is used in

Table 4.1: Details of slab models. Explanation of abbreviated labels is as follows. O: overlayer, M: mixed, S: sublayer, LC: low-coverage limit

Surface	Supercell	#O	Coverage / ML	#Layers / #Relaxed	KPOINTS
Ni(100)–O	p(1×1)	1	1	5L/3R	9×9×1
Ni(100)–M	p(2×1)	2	1	5L/3R	5×9×1
Ni(100)–S	p(1×1)	1	1	5L/3R	9×9×1
Ni(100)–LC	p(3×3)	1	1/12	5L/3R	3×3×1
Ni(111)–O	p(1×1)	1	1	5L/3R	9×9×1
Ni(111)–M	p(2×1)	2	1	5L/3R	5×9×1
Ni(111)–S	p(1×1)	1	1	5L/3R	9×9×1
Ni(111)–LC	p(3×3)	1	1/9	5L/3R	3×3×1
Ni(211)–O	p(1×1)	3	1	12L/8R	9×3×1
Ni(211)–M	p(1×1)	3	1	12L/8R	9×3×1
Ni(211)–S	p(1×1)	3	1	12L/8R	9×3×1

the energy calculation. The formate structure candidates are shown in Figure 7 (on the right from top to bottom: monodentate, bidentate-planar, bidentate-tetrahedral). The monodentate structure corresponds to the configuration in solid-state nickel formate dihydrate,[23] but it is not stable in the gas phase. The bidentate planar structure is the most stable one. Note that this is a 16-electron square planar structure, common for d8 metal complexes.

The reaction energy for the simpler 1ML case is given by the following equation.

$$\Delta G_{\mathrm{et}} = [G_{l+,\mathrm{NiO}} - G_{l,\mathrm{Ni}} + (n^{\circ}_{\mathrm{Ni}} G_{[\mathrm{Ni(HCOO)_2}]} + n_{\mathrm{Ni}} G_{\mathrm{H_2O}} - 2n^{\circ}_{\mathrm{Ni}} \cdot G_{\mathrm{HCOOH}})] \cdot \frac{1}{n^{\circ}_{M}} \qquad (4.3)$$

The derivation of the reaction energy for the low coverage limit using the layer-removal model is given by equations (3.5) and (3.6).

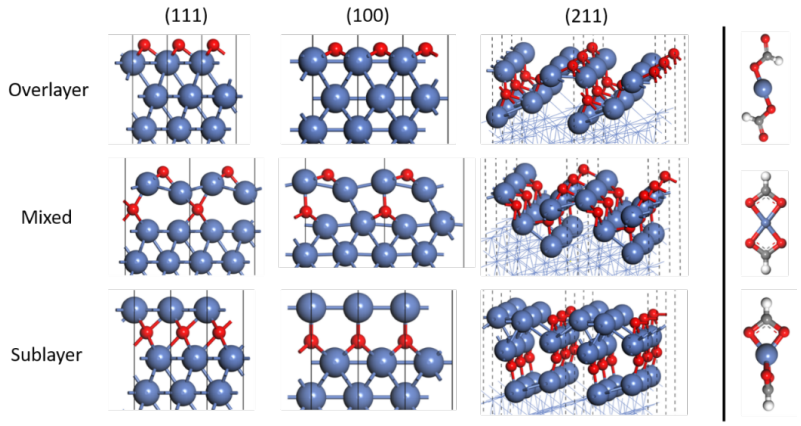


Figure 4.4: (a) Illustration of the relaxed structures of (100), (111), and (211) surfaces with (1) overlayer, (2) mixed, and (3) sublayer oxygen adsorption. The (111) and (100) surfaces are shown in side view, and the (211) surface is shown in perspective. (b) From top to bottom: monodentate, bidentate planar, bidentate tetrahedral geometry for the Ni diformate complex. The bidentate planar geometry is calculated to be the most stable of all, being 0.41 eV lower in energy than the tetrahedral stereoisomer, which is in turn 0.99 eV lower in energy than the monodentate structure.

4.4 Results

: Figure 4.5 shows the reaction energy of various activated surfaces considered (per Ni atom removed). Notably, for the 1ML case, oxygen-activated nickel surfaces can be favorably etched via formate chemistry, consistent with the experimental work. Etching of the pristine nickel surface to produce Ni diformate and gas-phase hydrogen is calculated to be endergonic by 1.7 eV and is hence highly unlikely. This is consistent with experiments where the unmodified clean nickel is not etched. Purple bars on figure 4.5 show that the etching reaction for Ni surfaces modified by a low coverage of oxygen adatoms is also thermodynamically unfavorable by more than 1.5 eV. Hence only surfaces modified by a high coverage of O atoms, with occupation of subsurface sites and the formation of a surface NiO layer, result in a favorable etching reaction with formic acid, in agreement with experiments. Note that the favorable etching reaction for these oxidized surfaces does not result from a greater stability of the products (Ni di-formate and water) since the energies are already normalized with respect to the number of metal atoms, but from a reduced binding of the Ni atoms in the "reactant" surface.

The site dependence trend is quite complex. For the (111) surface, increasing access to sublayer sites increasingly destabilizes the surface, while (100) and (211) surfaces are not significantly destabilized by occupying a fraction of the sublayer sites in a mixed adsorption structure. Results for (100) and (211) surfaces indicate that occupying the sublayer sites using low energy oxygen ions can make the etching reaction favorable. A closer look at figure 4.5 reveals that the etch becomes more favorable when Ni atoms in the top layer have a longer bonding distance with the layers underneath due to occupation by oxygen atoms in the sublayer. In particular, all the sublayer structures show a large vertical relaxation of the top layer, moving away from the bulk, while the mixed and overlayer structures still have one or more Ni atoms at a short bonding distance from the underlying metal. This suggests that the geometry distortion caused by the oxygen modifier provides a favorable driving force. However, a more detailed search of the configuration space is required to validate this. Overall, the (211) surface has a higher reaction energy compared to (100) and

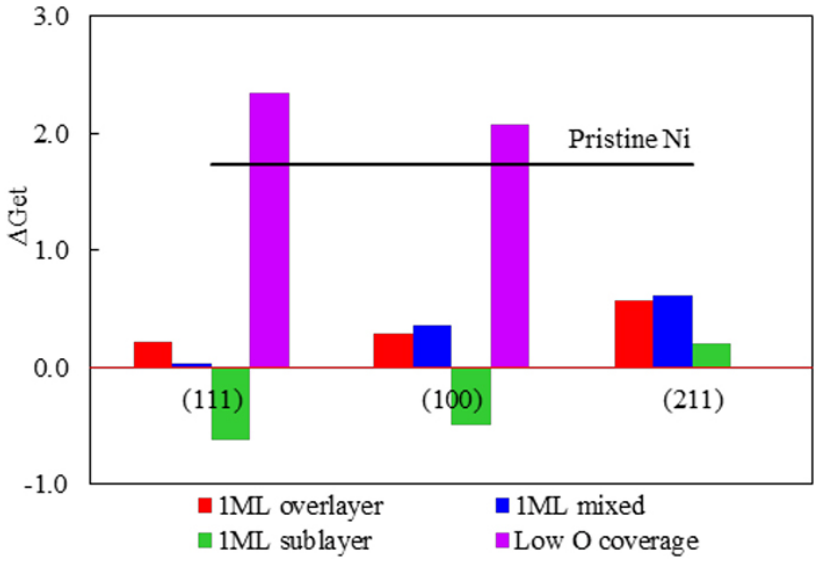


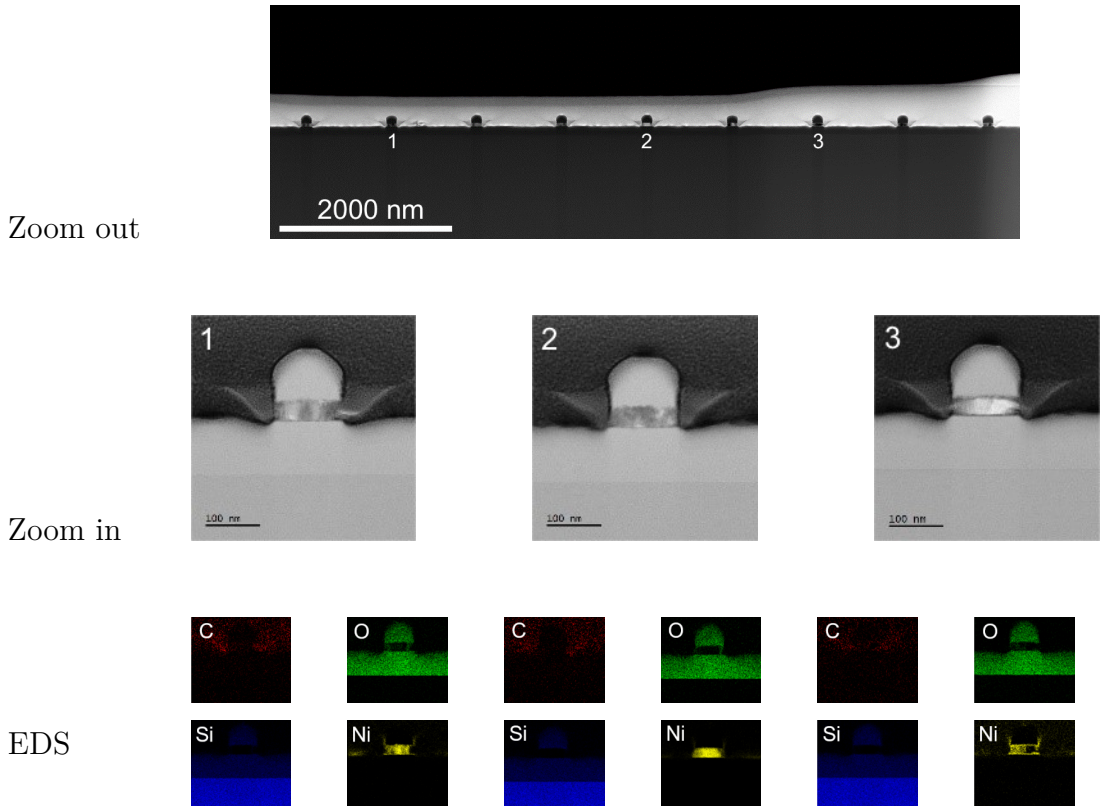
Figure 4.5: Etching reaction energies (ΔG_{et} in eV) of oxidized nickel (111), (100), and (211) surfaces using a layer-by-layer removal model (negative values indicating favorable etching, see supplementary material), normalized to one Ni atom removal. Low oxygen coverage does not result in a favorable etch, as indicated by positive reaction energies of over 2 eV. Sublayer sites significantly reduce the energy cost, in some cases resulting in a favorable etch, as observed experimentally. The energy level for a pristine nickel surface (1.74 eV, black line) is calculated with Ni bulk formation energy since the same surface is exposed if the entire atomic layer of nickel is etched and removed.

(111) surfaces. This is due to the fact that while the (211) surface itself has a higher surface energy, occupying the undercoordinated sites on the edge initially stabilizes the surface to a higher extent than that of (100) and (111) surfaces, where the terrace atoms have a higher coordination number. These results suggest that site specificity can play a role in atomic layer etching of surface species.

The feasibility of using low-energy oxygen ions from a plasma, as specified above, to convert a metal to a metal oxide surface without physical sputtering, followed by a formic acid reaction to form volatile reaction products, is therefore established. The effect of modification from low energy oxygen ions could be further leveraged for directionality, enabling anisotropic chemical activation in selected open areas of a pattern. An example shown in section 4.4 illustrates the effectiveness of such an ALE process. Starting with a 40 nm Ni film that was patterned with a hard mask of SiO_2 lines (SiO_2 thickness = 95 nm, line width = 125 nm, and line pitch = 1 µm, initial sidewall angles of 60° to 80°), low energy oxygen plasma was used to tailor the formation of nickel oxide, followed by nickel oxide removal with formic acid vapor exposure. The zoom-out TEM images confirm the process uniformity (over at least hundreds of nanometer), while the electron dispersion spectroscopy (EDS) mapping of Ni shows an 87° final sidewall angle. Ni was completely removed in all exposed region, while slight variations in the sidewall profiles were noted from elemental mapping by EDS. These variations may be attributed to the grain size and grain boundary effects, which are beyond the scope of this current work but an important aspect to be researched further.

To further enhance the control of surface oxide layer formation, an ideal approach would involve using a mono-energetic and reactive ion source. This would result in a narrower ion energy distribution, leading to a more precisely defined oxide layer and ultimately enabling greater precision in atomic layer etching. The intrinsic directionality of oxygen ions would also contribute to a directional formation of NiO in the presence of hard masks, resulting in an anisotropic removal of the oxides.

It is important to note that the results presented in this chapter are calculated for the nickel formate monomer product (see figure 4.6), while the product used in chapter 3 is the most stable nickel oligomer. To facilitate comparison, the relevant energetics for the results



in chapter 3 are reproduced in appendix A.6.

4.5 Conclusions

The results presented here highlight the effect of low-energy reactive ions and a low neutral-to-ion ratio on atomic layer etching. This opens up a new pathway to tailor surface reactions for selectivity and directionality. It is shown that the now much greater chemical design space is best explored using computational tools that can rapidly investigate selectivity and site-specificity. While the focus of this work is on oxidation-enabled nickel ALE, the principles explored are expected to be general. With the computational tools developed, it is possible to extend the process to less-studied materials with tighter dimension constraints and reduce the time from lab-scale research to commercial-scale processes. Plasma-thermal ALE of metals is motivated and discussed, in combination with first-principles calculations and experimental confirmation, to demonstrate how surface reactivity and selectivity, rather than the etch rate, are the focus of realizing nanometer-level patterning of etch-resistant

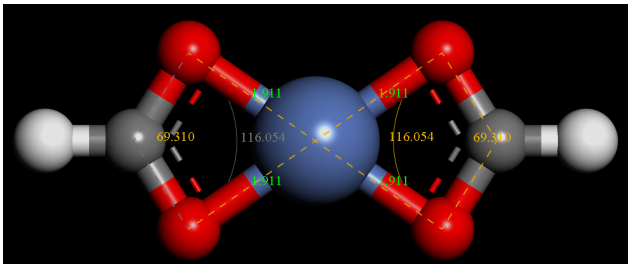


Figure 4.6: The optimized structure of nickel formate complex in bidentate square planar configuration. The bond lengths and bond angles are marked in degrees and angstroms, respectively. The two ligands are completely in plane with no distortion.

materials.

In the next chapters, armed with machine learning tools that can explicitly probe many of these effects, the problem of controlling ion energy and ion-to-neutral ratio are revisited with atomic resolution to complement the XPS results presented here.

CHAPTER 5

Neural Network Potential for Cu surface - O_2 Plasma Interaction

5.1 Introduction

In this chapter we discuss the development of an machine-learning interatomic potential (MLIP) to describe the interaction between the copper surface and the oxygen plasma. Particular attention is paid to the practical aspects of MLIP development regarding training, inference efficiency, dataset gathering, and accuracy. The chapter is divided as follows. In section 5.1 an motivation of the interatomic potential is given and the state of the art is reviewed regarding both the methodology and the results. The computational details are given in section 5.3.5. In section 5.3 a detailed discussion is given regarding how the dataset is constructed. This process is presented illustrating the logic involved, as the problems encountered here are general and the solutions applicable to other systems. In section 5.4, chemically-intuitive validation results were shown. We discuss and quantify some inherent errors involved in our approach in section 5.5 and conclude in section 5.6.

To the author's knowledge, there have been no molecular dynamics studies on the interaction between a metal surface and the nonthermal oxygen plasma. Here, "nonthermal plasma" refers to the fact that in the plasma used for material processing, the electronic temperature is typically much hotter than the temperature of the heavy species (ions and neutrals). Plasmas are sustained by scattering of accelerated electrons with heavy species, causing excitations and ionizations. Previous surface-plasma interactions studies usually involve Si-based materials,[32,70,80,85,136] and are performed with classical interatomic potentials (*e.g.* Tersorff-Brenner[17,18] & Stillinger-Weber[15]). These simulations set the ground for

computational studies of surface-plasma interaction at the atomic scale, but they are fundamentally limited in accuracy by the functional forms of the potentials. Roughly speaking, there are two families of classical interatomic potentials, based respectively on the embedded atom method (EAM,[13] MEAM,[22] MEAM-Qeq[21]) and the bond order method (COMB,[87,96] ReaxFF[52,120]). These formalisms are designed with metallic and covalent systems in mind, respectively, and there is little knowledge on their accuracy when the same set of parameters is required to reproduce the metallic (in the pristine layers), the covalent (in the plasma, in the physisorbed molecules, and to some extent the oxide), and the ionic bonds (in the oxide) interactions at the same time. Recently ReaxFF[90] and COMB[89,96] have been applied successfully to the thermal oxidation of copper, but their functional forms prevent the accurate description of the regions of the potential energy surfaces encountered during plasma oxidation, which could involve structures with very short interatomic distances.

The development of machine learning potentials such as the Behler-Parinello high-dimensional neural network potentials (HDNNP)[68] and the Gaussian approximation potentials (GAP)[82] provide functional forms that are versatile enough to be parametrized to describe highly diverse chemical environments at or near the accuracy, previously only achievable with expensive first-principles calculations.

To provide a comprehensive description of the plasma, it is crucial to accurately account for the various components, including charged ions and neutrals in their ground and excited states (both electronically and vibrationally), free electrons, fast ions, and radiation. However, in existing explicit atomistic treatments of plasma-surface interactions known to the author, most components except for fast ions are often disregarded. Additionally, the ions are typically treated as fast neutrals, assuming that the neutralizing species are the secondary electrons repelled by the sheath. In this work, we adopt the same approach.

Most machine-learning models face challenges in effectively handling long-range Coulomb interactions. Moreover, since we have adopted plane-wave Density Functional Theory (DFT) as the reference method, the treatment of charged systems necessitates implicit solvation methods and charge balancing, significantly increasing computational costs. Even with these considerations, the ambiguity in charge partitioning schemes in plane-wave DFT introduces

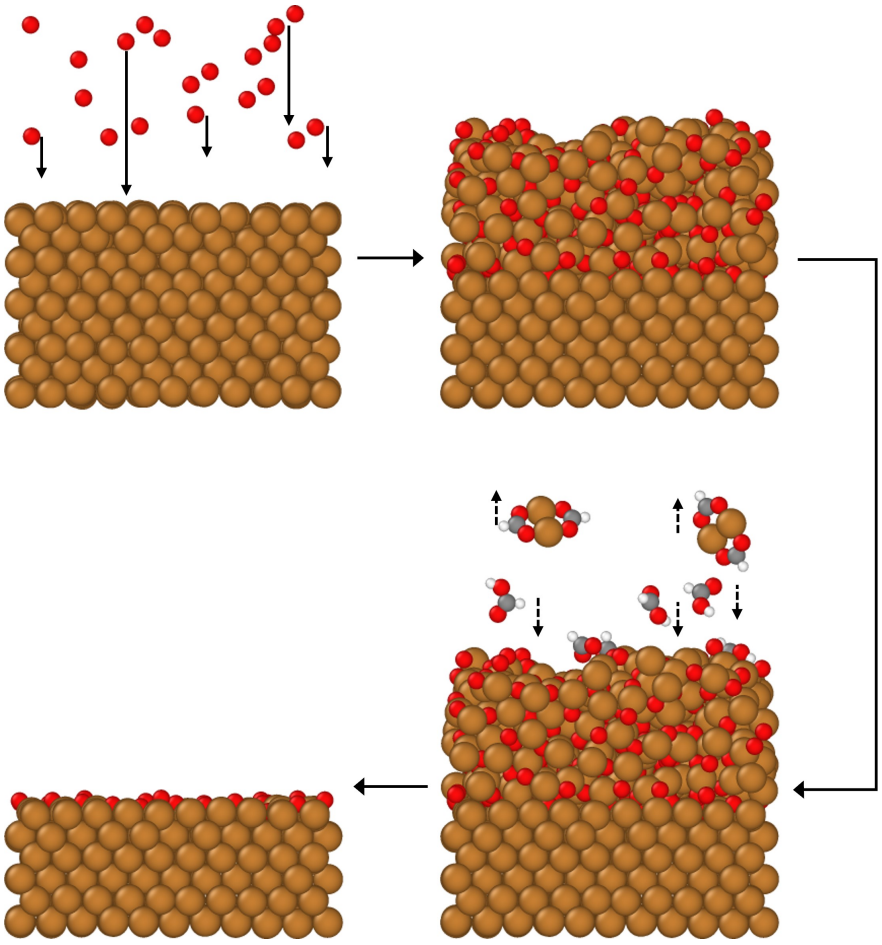


Figure 5.1: Overall process of plasma-thermal ALE on Cu metal. **a)**: Starting with a pristine surface, reactive oxygen neutrals and ions are generated by the plasma. The ions are accelerated by the plasma sheath toward the surface, indicated by the long arrows. The neutrals, much larger in number, diffuse toward the surface with thermal motion. **b)**: The top layers of the surface is converted to a copper oxide. **c)**: Plasma is purged, and etchant molecules introduced to the chamber. Etchant reacts with the surface oxide, producing volatile organometallic complexes that diffuse away. **d)**: etch stops when the oxide is consumed.

further complexity.

Regarding electronically excited states, explicit time-dependent DFT (TDDFT) or its linear-response variant is required, but both approaches are still computationally expensive for dataset generation. On the other hand, vibrationally excited states are adequately accounted for on the surfaces. The ions utilized in our study are atomic ions and do not possess vibrational degrees of freedom.

In this chapter, it is shown that high-dimensional neural network potential (HDNNP) is capable of describing all stages of copper oxidation at an accuracy close to that of the underlying DFT method and hence can be used as a reliable alternative to existing reactive force fields. Practical limitations of the current machine-learning approaches are highlighted and future directions of improvement are specified.

5.2 Theoretical Methods

All electronic structure calculations are performed with the density functional theory as implemented in the Vienna ab initio Simulation Package (VASP).[27,34,35] The electron-ion interactions are treated using the projector augmented wave (PAW) method[48] and the valence one-electron functions are developed on a basis set of plane waves. The PBE exchange-correlation functional[36] is used throughout. While it is true that hybrid-functional generally yield better descriptions of metal oxides, it is inapplicable due to the computational cost (see section 5.2.1) and its poor description of the pristine metal. The bulk crystal parameters are obtained by a two-step direct volume relaxation starting from their experimental values.[115] For all systems, the cutoff energy of the plane wave basis is set to 460 eV. The energies are converged to 10^{-6} eV. For the metal slabs in the training set calculations, second-order Methfessel-Paxton smearing was used with a width of 0.2 eV. For the oxides, Gaussian smearing with a width of 0.02 eV was used. All calculations were done spin-polarized. Magnetism in most cases is small ($< 3\mu_B$ for the (3×3) slab) but not negligible.

5.2.1 Active learning and delta learning at hybrid level of theory

The number of required structures was estimated to determine the feasibility of training a neural network potential at a higher level of theory (e.g. hybrid functionals), following the automatic training method proposed in Ref[163]. This method estimates the minimum number of structures required to train a neural network potential. The method is outlined as follows:

1. From a large pool of structures labelled with energies and forces at the PBE level, randomly select 10 structures.

2. Train two neural network parametrizations `NNP-1` and `NNP-2`, differing by the random seeds of their starting points.

3. Evaluate energies and forces on all the structures in the dataset, using both potentials.

4. Select the 10 structures where `NNP-1` and `NNP-2` disagree the most in terms of energy, and add to the training set.

5. Retrain both parametrizations. This constitutes one round of the automated training process.

6. After each round, the RMSE between `NNP-1` and `NNP-2`, between `NNP-1` and `DFT`, and between `NNP-2` and `DFT` are recorded.

figure 5.2a shows that at 150 structures, the validation set error cannot be minimized to a satisfactory value. Note that the unit of the error is in eV per atom as opposed to meV per atom. For all data points shown, the error in the training set is less than 5 meV per atom. Training with such a small training data set is prone to overfitting, hence the learning curve is monitored to ensure error on the validation set (the rest of the pool) does not increase with increasing epochs. Since the disagreement between the two networks is much smaller than those between the networks and the ground truth, we clearly do not have enough data in the training set.

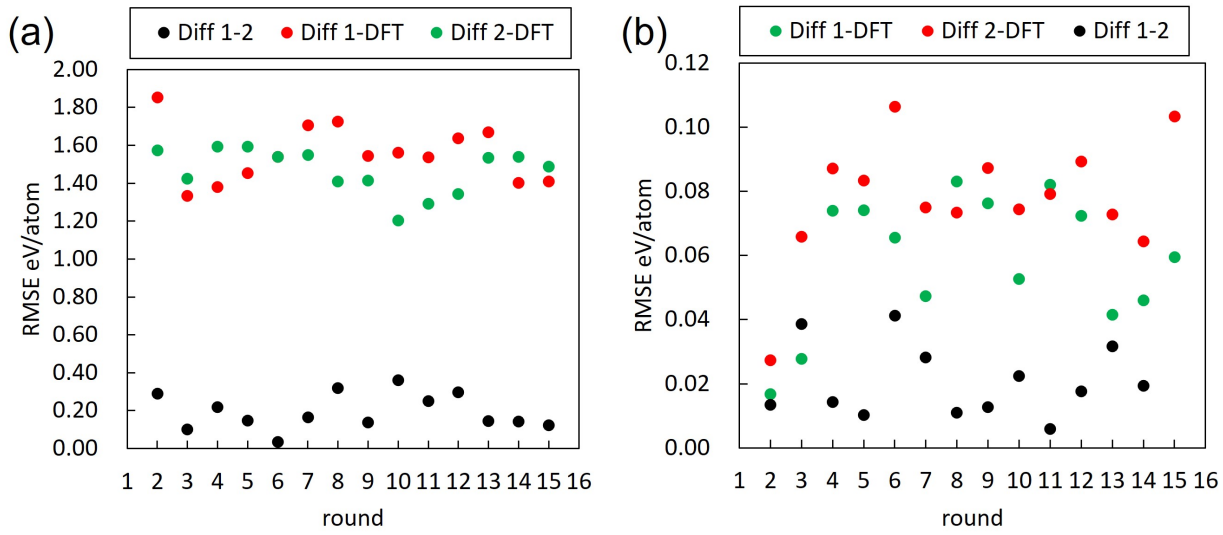


Figure 5.2: **a**: validation errors for the automated training method. "diff 1-2" represents the difference between the two network potentials. "diff 1-DFT" and "diff 2-DFT" represent the difference between the network potentials and the DFT ground truth. All data points are evaluated over the whole pool of structures.

Additionally, attempt was made to use the delta learning approach, which has been shown to require a much smaller dataset. The corresponding result is shown in figure 5.2b. To save on computational cost in this proof-of-concept investigation, the network potential is used to fit the differences between the PBE and RPBE exchange-correlation functionals, as opposed to a high-level method. Note that the error is decreased by a factor of almost 10 compared with the brute-force fitting to the PBE potential energy surface, indicating that a smaller training set may be used. However, at 150 structures (15 rounds), we do not see a significant decrease in the NNP-DFT difference.

Given the high cost of hybrid functional calculations, a training set much larger than 150 structures cannot be afforded. Therefore, we conclude that from a practical point of view, we can only use the GGA-level of theory. We also note in passing that the exchange-correlation functional must describe properties of both the metal, the molecular and atomic modifier, and the oxide phases accurately. While hybrids could improve the description of the correlated oxide, they in general do not describe the metals well. Hence, the versatile PBE functional is used.

The molecular dynamics simulations were performed using the Large-scale Atomic / Molecular Massively Parallel Simulator (LAMMPS).,[175] in conjunction with the ReaxFF package[94] and the NNP interface to the n2p2 neural network potential library..[146] The standalone ReaxFF package was used to reparameterize the inner-wall parameters to stabilize the very initial ReaxFF simulations.

In this contribution the high-dimensional neural network potential (HDNNP) of Behler and Parinello was chosen as the machine-learning framework. This is motivated by theoretical and practical reasons.

In chronological order, the first potential attempted was the ReaxFF potential. However, it was difficult to fit the ReaxFF functional form due to three reasons. The first two reason is physical: on the one hand, the functional form is inspired by physics and in diverse environments it is difficult to know a-priori which contributions to the energy should be fitted, and how the dataset should be prepared. On the other hand, when using ReaxFF for MD simulation, it is impossible to know if the potential explored regions not fitted well by the hand-crafted training dataset. Arguably, the protection against extrapolation in HDNNP is limited as well, but better than nothing. The third reason is practical: the ReaxFF fitting program works by optimizing one parameter at a time by probing the loss function (sum of reaction energy errors) at three points using a 3-point sampling around the current parametrization. From these three points a parabola was constructed and the parameter value that give the minimum of this parabola is chosen as the next parameter value. This rudimentary method clearly is not suitable for large datasets, and in practice we found this optimization method highly sensitive to the ordering of optimized parameters, since the code works on one parameter at a time. Note that the situation has improved significantly with tools constructed since this work.

The high-dimensional neural network potential formalism was chosen mostly because of practical constraints. While the rapid advancement of machine learning techniques have improved the situation dramatically, at the time of project onset, the n2p2 package was the only package offering very fast ($\sim 5\,\mu s$ per atom per core) inference speed through intelligent intermediate result re-use in the featurization step and stable, efficient LAMMPS interface,[146]

very fast training convergence with the Kalman filter method.[147] These practical constraints lead to `n2p2` being chosen.

In terms of the neural network potential architecture, the atomic neural networks of Cu and O consist of 85 nodes in the input layer and 2 hidden layers with 30 nodes each. Neural network potential training was carried out using n2p2's training routines (`nnp-train`)..[147] The fingerprints were chosen to be Behler-Parinello style symmetry functions, with the η, r_s, λ, ζ parameters generated automatically according to Ref.[138]. The fingerprints are then "pruned" based on the variance of their values across the entire dataset: fingerprints whose value stays constant do not capture the variations in chemical environment and are removed. The n2p2 input file `input.nn` is given in listing B.1. The architecture, the symmetry function parameters, etc. can be viewed directly from this file.

5.2.2 The physics-informed potential: ZBL at short distances

To readily capture the strong, exponentially growing potential energies at extremely short distances that occur during ion impact and ensure the ML-enabled simulation is stable, the neural network potential formalism has been revised to include a physics-inspired functional form at very short distances. Specifically, an "inner cutoff" radius r_c, is introduced the smoothly suppress the energy result from the fitted neural network over a range of $r \in (1, r_c)$. This suppression is performed using a function of the shifted cosine form:

$$f(x) = \frac{1}{2}[-\cos(2\pi x) + 1]$$
$$x = \frac{r - r_1}{r_c - r_1}$$

The updated cutoff function with both the "inner cutoff" introduced here and the original neighborlist "outer cutoff" $r_{c,nn}$ are shown in figure 5.3. With the underestimated, unreliable short-distance repulsion from NNP removed, a strong repulsion must be added. For this purpose, a screened Coulomb potential of the form proposed by Ziegler-Biersack-Littmark

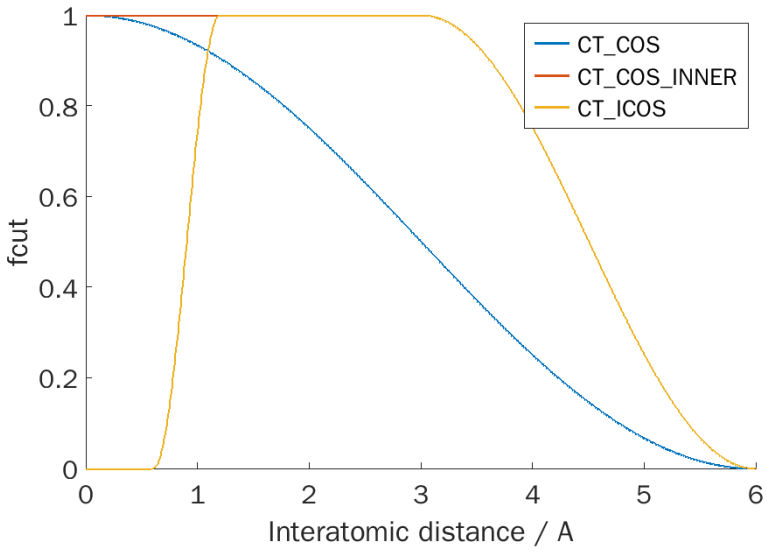


Figure 5.3: The inner cutoff function introduced for smoothly suppressing the NNP energy contribution at short distances

(ZBL) is introduced for the bond distances $r < r_i$. The ZBL functional form is given below:

$$
\begin{aligned}
E_{ij}^{ZBL} &= \frac{1}{4\pi\epsilon_0}\frac{Z_i Z_j\, e^2}{r_{ij}}\phi(r_{ij}/a) + S(r_{ij}) \\
a &= \frac{0.46850}{Z_i^{0.23} + Z_j^{0.23}} \\
\phi(x) &= 0.18175e^{-3.19980x} + 0.50986e^{-0.94229x} + \\
&\quad 0.28022e^{-0.40290x} + 0.02817e^{-0.20162x} \\
S(r) &= C \qquad r < r_1 \\
S(r) &= \frac{A}{3}(r - r_1)^3 + \frac{B}{4}(r - r_1)^4 + C \qquad r_1 < r < r_c \\
A &= (-3E'(r_c) + (r_c - r_1)E''(r_c))/(r_c - r_1)^2 \\
B &= (2E'(r_c) - (r_c - r_1)E''(r_c))/(r_c - r_1)^3 \\
C &= -E(r_c) + \frac{1}{2}(r_c - r_1)E'(r_c) - \frac{1}{12}(r_c - r_1)^2 E''(r_c)
\end{aligned}
$$

It can be easily verified the switching function $S(r)$ ensures the potential is smooth and differentiable to the second order at the inner cutoffs r_1 and r_c. The parameters r_1 and r_c

are pairwise-specific and are fitted to the Cu−Cu, Cu−O, and O−O dimer potential energy surfaces. Two comments are due here. First, this modification does not change the potential

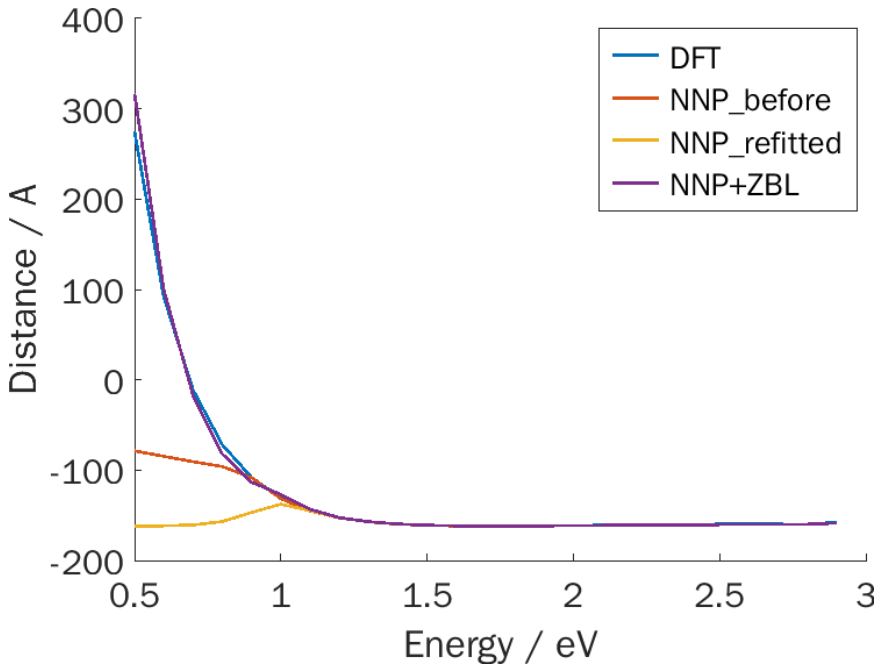


Figure 5.4: Comparison of fitted 0 eV to 400 eV ZBL+NNP potential energy surface with DFT

energy surface within 20 eV from the minima. Therefore the validation results are unaffected. Second, the usage of a pairwise interaction is motivated by the fact that it is unlikely for an atom to be simultaneously close to multiple neighbors, as it would require these neighbors to be close themselves before the impact. Since the energies considered here far exceed those involved in chemical bonding, any near-equilibrium structures, as found on the ion impact targets, would not have such structures. Therefore, using a pairwise potential should not incur significant error compared to a many-body description. Third, to perform simulation using this potential, a patch to the n2p2 package is available from its public repository. An extra statement `inner_cutoff` f_1, f_c in the `input.nn` is required to switch on this smooth inner-wall suppression of the NNP. The parameters $f_1 = \frac{r_1}{r_{c,nn}}$ and $f_c = \frac{r_c}{r_{c,nn}}$ refer to the ratios of the newly introduced inner cutoff parameters with to the outer, neighborlist cutoff distance. Then, within LAMMPS, the `pair_style hybrid/overlay` command is used to combine the ZBL and the NNP potentials.

5.3 Dataset generation

There are several types of structures used in the training set generation and production MD simulations. Throughout this paper, the convension ($w \times l \times t$) is used in denoting the dimension of the slab, where w, l, and t refer to the width, the length, and the thickness of the slab structure with reference to the conventional Cu(100) primitive surface cell (1 atom, single layer). In training set generation, the majority of the structures started from a Cu(100) ($3 \times 3 \times 5$) surface which is subsequently oxidized. A small subset of training MD simulation started with ($3 \times 3 \times 10$) Cu(100) surfaces. For these surfaces, a ($3 \times 3 \times 1$) k-space mesh was used. Furthermore, bulk oxide structures Cu_2O and CuO were used with ($2 \times 2 \times 2$) supercells of conventional unit cell for both. The reciprocal space was sampled with ($3 \times 3 \times 3$) and ($5 \times 5 \times 3$) meshes for Cu_2O and CuO, respectively. Additionally, a ($8\sqrt{3} \times 8\sqrt{3}$) supercell for the Cu(111) surface was used in appendix C.3. In these simulations, the number of layers is not of interest because the oxide never exhausted the copper layers.

5.3.1 MD sampling procedure

To train a neural network that is able to describe all the stages of oxidation (chemisorption, surface oxide growth, and bulk oxide growth), an iterative approach based on molecular dynamics was adopted. This procedure requires a starting energy engine to generate the very first dataset. For the present study, a ReaxFF parameterization reported in the literature[165] was chosen. A slight reparametrization of the inner-wall repulsion terms was necessary to stabilize the training set generation MD simulations, as the original potential was used for thermal surface catalysis conditions.

The starting point of the parametrization was taken to be the ReaxFF potential. Since original training set does not include short-distance interactions of Cu/O, the as-found potential is not very stable to plasma impact even at a kinetic energy of 5 eV. As a result, the inner wall repulsion term, implemented in ReaxFF, was re-parametrized to the potential energy curves of Cu-Cu, Cu-O, and O-O dimers, calculated at the PBE level. The comparison of refitted O_2 diatomic molecular potential energy surface is compared with DFT values in

figure 5.5. This step ensures the exponential inner wall repulsion is reproduced.

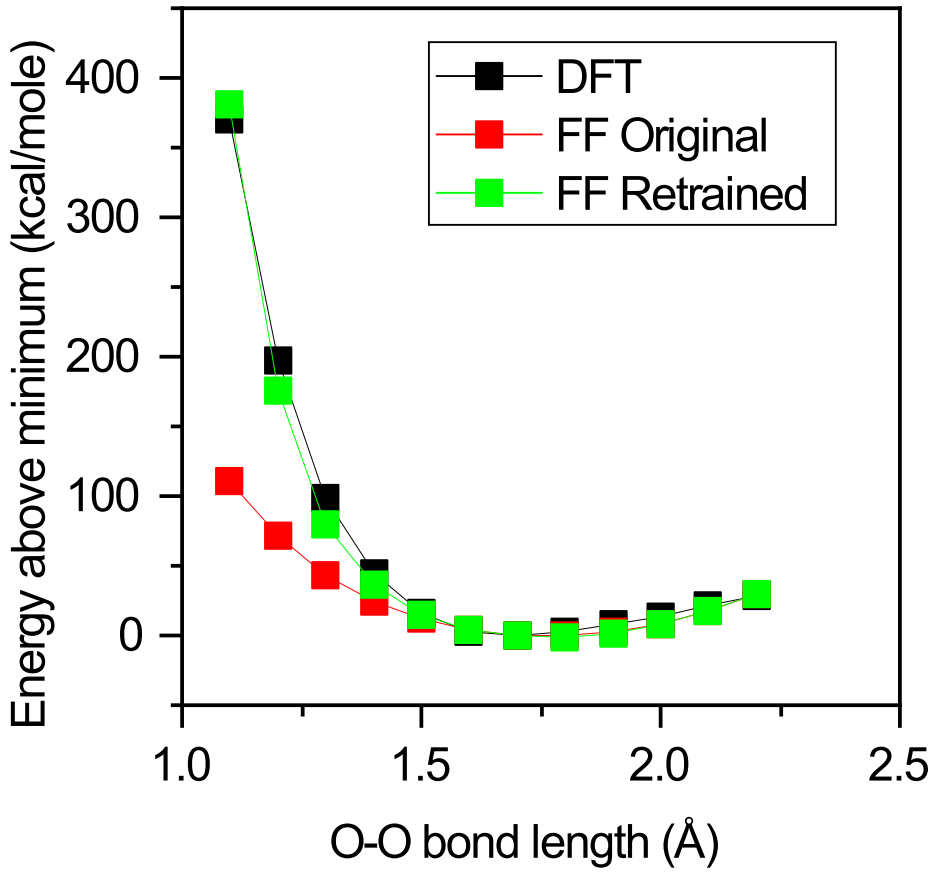


Figure 5.5: comparison of potential energy of a O_2 molecule using literature ReaxFF parameter set (FF original), reparametrized ReaxFF parameter set (FF-retrained), and DFT.

After MD_1 was generated, ReaxFF is no longer used. This leads to the dataset MD_0, used to train the first-generation NNP_1. After this point, the approach is summarized in the following steps:

1. Starting with potential NNP_i, trained using MD_i and possibly earlier datasets.

2. Perform MD simulations for training set generation using NNP_i, using setup similar to production MD, but on small substrates efficient for DFT.

3. Using DFT, calculate the ground state energies and forces on the snapshots extracted from the trajectories. This yields MD_{i+1}. In this case, ~ 1000 structures was included every iteration. This number is not optimized and is likely to be more than what is necessary.

4. If NNP_i gives low errors ($< 5\,\text{meV/atom}$) relative to DFT, stop. Else, add MD_{i+1} to the dataset, or replace the whole dataset with MD_{i+1}. Whether the new dataset is used to (partially) compound or replace the existing dataset is mainly motivated by the need to keep the size of the training set manageable in terms of training time. Note that often it is reasonable to replace the old datasets with the new ones because the PES covered by the new dataset encompasses that of the old one, in addition to better sampling the configurations reachable in the DFT PES.

5. Using the new dataset, retrain to obtain NNP_{i+1}, and repeat.

Note that, while the training procedure is applicable to different substrate/modifier pairs, a usable interaction potential may not always be available. In that case, this step can be replaced with DFT-driven *ab initio* molecular dynamics at crude computational accuracy, or a semi-empirical method.

The training set generation MD involves oxygen species (O_2 at early iterations, O atoms at late iterations) impacting Cu(100) (3×3) surfaces. The snapshots were sampled to focus on the impact events (see section 5.3.2). Data cleaning is done before and after DFT calculations, based on geometric and energetic criteria, respectively (see section 5.3.2).

In the procedure described above, when the neural network potential is deemed to perform similarly well on the predicted snapshots as on the training snapshots, the training is stopped. At this point, the kinetic energy E_k increases or the molecular species are replaced with the atomic species. To cover the 0 eV to 20 eV kinetic energy range, the projectile is gradually changed from 10 eV molecular, to 10 eV atomic, to 20 eV molecular, to 20 eV atomic. Given

that the bond energy of O_2 is 5.16 eV, using atomic projectile is equivalent to molecular projectiles with ~ 5 eV higher kinetic energy. Thus, this method is equivalent to ramping up the kinetic energy in steps of 5 eV, while keeping the network aware of both O_2 and O interactions. This can be considered as an outer training loop on an energy scale. The starting point in all these trajectories are $(3 \times 3 \times 5)$ Cu (100) slabs. These datasets are simply labelled `md`. The details are listed below: As can be seen from the table, the size of

Table 5.1: Details of the `md` part of the dataset

label	ion type	E_k	number of configurations	comments
`md0`	molecular	10 eV	5397	ReaxFF
`md1`	molecular	10 eV	741	
`md2`	atomic	10 eV	1482	
`md3`	molecular	20 eV	2565	
`md4`	atomic	20 eV	424	
`md5`	molecular & atomic	20 eV	206	resampled `md3` and `md4`

the training set at each step is highly different, due to the attempt to (non-systematically) minimize the number of DFT calculations. At each of effective E_k, much more structures are generated than fed to DFT, which was performed in small batches until the prediction error is lowered. This procedure is effective but tedious, and effort is being made to automate the process. Below the details are given for the first iteration.

The training routines were able to minimize the energy and force errors on the test set to 2.32 meV/atom and 0.15 eV/Å, respectively. The new neural network parametrization, termed NNP_1, was used in identical MD impact simulations as those used to generate `md0`, to generate a new set of snapshots. Again, 25 distinct MD runs were performed. After a cleaning procedure identical to that applied to `md0`, yielding `md1` dataset, containing 2716 structures. Before training was performed on this new dataset, the self-consistency of the neural network potential nnp_1 was tested by calculating its prediction error on the dataset obtained via MD driven by itself (`md1`), as an benchmark for the out-of-sample error. Not

surprisingly, the energy error is 88.67 meV/atom and the force error is 1.98 eV/Å. The fact that the out-of-sample error is almost two orders of magnitude higher than the in-sample error is used as an indication that the sample (ReaxFF-generated at this stage) has a different underlying "distritbuion" than the ground truth (in this case, the DFT PES), in accordance with statistical learning theory. Therefore, the neural network potential is retrained on the dataset composed of `md1`, to yield a test set error 3.14 meV/atom and 0.17 eV/Å, respectively on energies and forces. New MD structures were generated (`md2`), and again the out-of-sample prediction error is calculated to be 2.78 meV/atom and 0.13 eV/Å, respectively. This indicates that the neural network potential, within the MD setup used to generate the training configurations, does not extrapolate. As a finishing step, `md1` and `md2` are combined to retrain the neural network potential, yielding nnp_2. Thus, the inner loop is exited, the kinetic energy increased, and sampling is performed again.

A similar process is applied for subsequent iterations. As noted from figure 5.6, the `md0` and `md1` datasets are not used in the training in the final neural network potential. The reason for `md0` is that it is drawn from an underlying probability distribution of structures according to the ReaxFF PES. Hence, including it is detrimental to fitting the network to DFT. `md1` is excluded because it is likely subsequent datasets can cover the potential energy surface explored by it. Including it only makes the training time longer. `md5` also deserves some comment. It is a resampling from the trajectories in `md3` and `md4`, but including structures where a short Cu-O bond is present. Such structures was excluded by data cleaning in other `md` datasets as discussed above, but here the criteria is relaxed slightly.

5.3.2 Heuristics for snapshot extraction

While the addition rate of ions is fixed at one per 10000 steps, the ions are added at varying heights and the surface is not smooth. Therefore, to sample short-distance structures efficiently, an analysis program tracks the minimal distance between the O atoms and Cu atoms at each time step in the trajectory. This signal is inverted and peak analysis is performed. The peaks are treated as steps where collision events occurred and the steps near

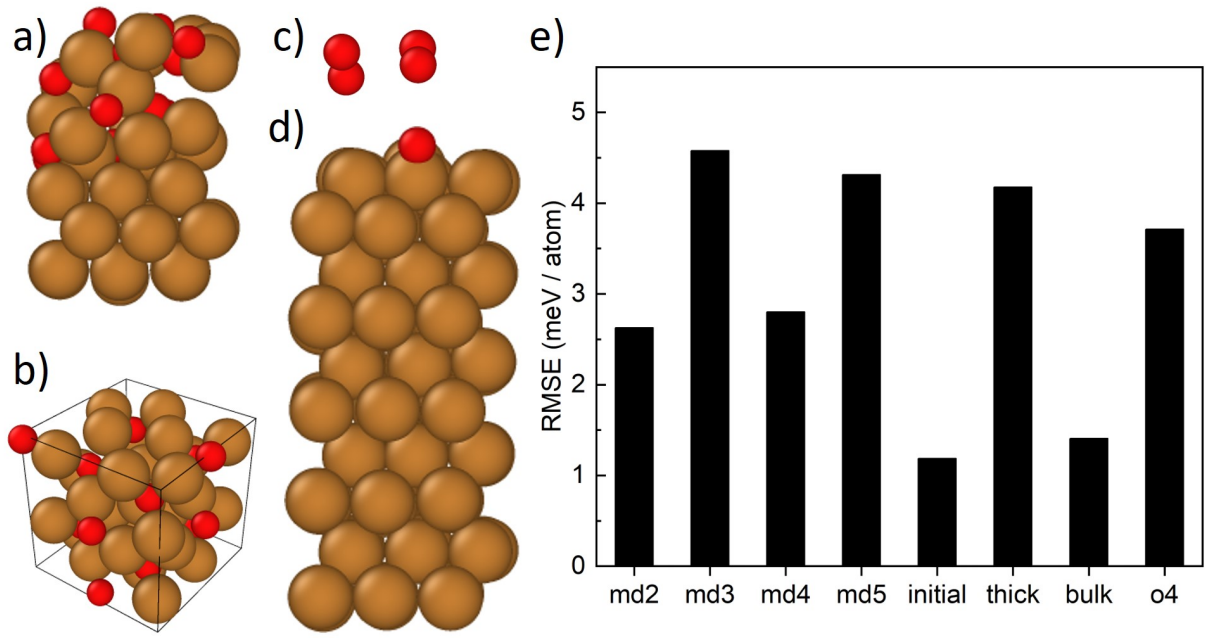


Figure 5.6: **(a)-(d)**: structures used in training: **(a)**: snapshots of Cu (100) slab under oxygen bombardment MD, this comprises the majority of the dataset; **(b)**: bulk copper oxide structure; **(c)**: O_2-O_2 interaction for oxygen clustering correction calculated by enforcing high-spin state; **(d)**: thick slab structures to minimize extrapolation error. **(e)**: errors on each subdivision of the dataset showing the potential performs well on all relevant environments.

the peaks are sampled at regular spacing. A collection of scripts for dataset generation and data analysis used in this study are available at GitHub.

It was found during training, that data cleaning is crucial to the success of training. Due to deficiencies in the early iterations of the network potential, the training MD often leads to unphysical atomic configurations that are few in number but problematic for DFT calculations. These structures cause the DFT calculations to either fail to reach self-consistency or give very high energies. The training routine cannot properly account for these scarce outliers, leading to persistent high errors during training. Therefore, data cleaning is performed to remove configurations with extremely short bond lengths. The cutoff values of the bond length depends on the target kinetic energy, and is determined by the corresponding bond length on figure 5.9. Additionally, in the very early stages, clearly impossible structures with

lost Cu atoms, disintegrated slabs, etc. are removed.

5.3.3 O_2 clustering problem

During testing simulations using early parameterization, it was observed that oxygen molecules occasionally form large adsorbed clusters on the oxide surface. The root cause of these clearly unphysical phenomena is found to be some incorrect behavior of the potential energy surface at short O_2-O_2 separation from the underlying electronic structure method. Figure 5.7 shows the potential energy surfaces of quintet ($^3\Sigma_g^- +^3 \Sigma_g^-$) and singlet ($^1\Delta_g +^1 \Delta_g$) dimers. At typical van der Waals distances ($d >\sim 3.5\,$Å), the quintet is $\sim 2\,$eV more stable than the singlet, in accordance with the well-known singlet-triplet energy difference of O_2.[59] The adjacent triplet molecules adopt an antiferromagnetic (AFM) coupling to form an overall singlet ground state complex, qualitatively consistent with the literature.[76] At $d < 2.3\,$Å, however, the PBE functional erroneously gives a stabilization of the singlet state on each O_2 molecule, ($^1\Delta_g +^1 \Delta_g$), resulting in a strong attractive interaction at too short distances compared with experimental intermolecular potentials and high-level multireference calculations.[44,56,111] On such structures, DFT electronic relaxation adopts the unphysical singlet configuration($^1\Delta_g +^1 \Delta_g$), which is only 0.7 eV above the van der Waals distance minimum($^3\Sigma_g^- +^3 \Sigma_g^-$), and hence is easily reached in impact dynamics simulations if the DFT-based "ground state" PES is followed blindly. A better description can be obtained by forcing the O_2 molecules to maintain a triplet state, as in the most stable configuration at van der Waals separation distance. Note that the erroneous and strong binding state is unrelated to the well-known true and weak binding state (at $d > 3\,$Å, regardless of spin state), which is captured by PBE and high-level wavefunction methods alike.[46,98]

To fix the clustering problem, additional cleaning was performed to automatically remove problematic configurations with inter-O_2 distances smaller than 2.2 Å from the training set. The "correct" dimer interaction is added back to the dataset by performing new O_2-O_2 interaction MD simulations (data set `o4`). In these simulations, two O_2 molecules are placed in a large simulation box and attached to each other by a spring. High-temperature MD is

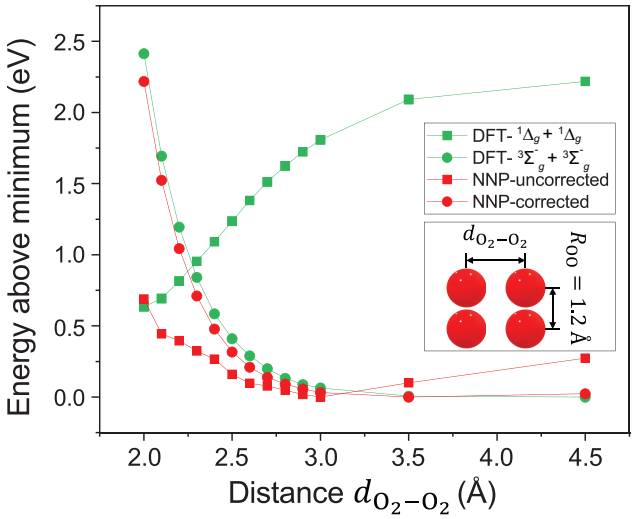


Figure 5.7: Comparison of the potential energy surfaces of O_2 complex in the "H" configuration, shown in inset. The corrected NNP is shown to follow the triplet state PES.

run to sample a total of 1990 configurations, which were subsequently calculated with DFT and forcing the high-spin state. Including this augmentational dataset leads to DFT PES being closely followed by the NNP (see figure 5.5).

The solution here (to the first problem) effectively used an overwhelming number of data points forced to have the correct orbital occupations to "convince" the model that the few lower energy points remaining in the dataset are outliers, and "drag" the model back to give the correct description. The successful application of this approach relied on two aspects: 1) the poorly described part of the potential energy surface can be isolated and is explainable using chemical reasoning. In other words, the deficiency in the underlying electronic structure method is clearly identifiable and the design of the correction data set (molecular O_2-O_2 interaction) is possible; 2) the correction dataset is much less computationally intensive than the problematic dataset. For situations with a larger chemical search space, both may become impossible. The error introduced by fluctuating orbital/band occupations when the underlying electronic structure method is applied to a large, diverse set of systems could become a major obstacle in parameterizing machine learning potentials.

In addition to the `o4` sub-dataset, other sub-datasets are used to correct for technical problems such as slight extrapolation when the cell size increased. The overall energy error

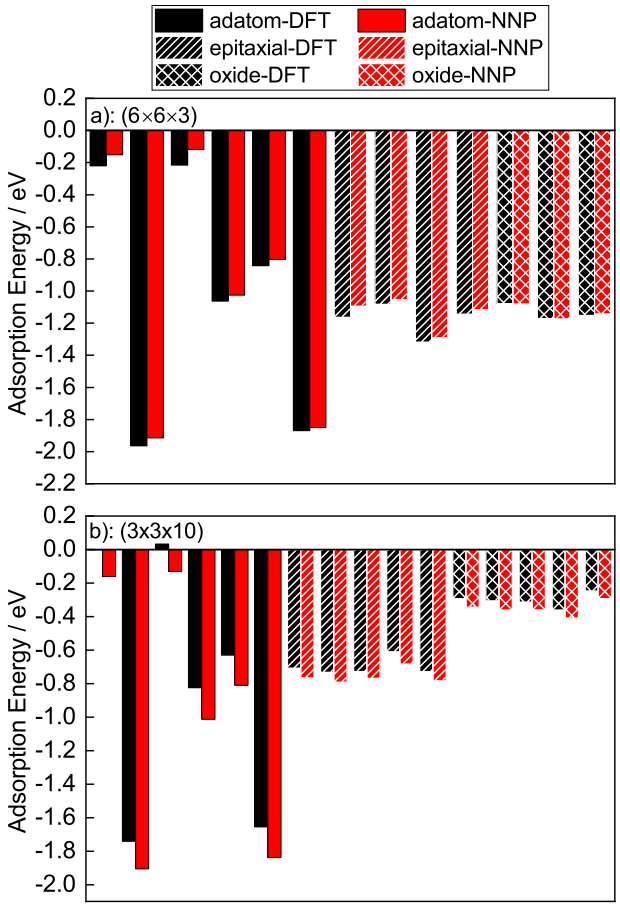


Figure 5.8: Comparison of average adsorption energies on realistic test cases including single adatom adsorption, epitaxial oxide on metal, and fully oxidized slabs. **(a)**: result on a $(6 \times 6 \times 3)$, thin Cu(100) slab. **(b)**: result on a $(3 \times 3 \times 10)$ Cu(100) slab.

of the resulting network is 4.17 meV/atom, and the overall force error is 0.11 eV/Å. The accuracy on each sub-dataset is shown in figure 5.6. It can be seen that all of the datasets have similarly low energy errors.

5.3.4 Low coverage problem

To prevent network extrapolation when moving from the small (3×3) training structure to the big (20×20) production structures (see chapter 6), additional DFT calculations were performed to sample better the initial impact where few oxygen atoms were present on the surface. Especially problematic was the case of a single O atom. Since the cell used in training MD is smaller in the lateral direction than twice the cutoff radius (12 Bohr) of the potential, the neighbor list of oxygen atoms likely includes other oxygen atoms if the are slightly separated from the center atom. Therefore, the O-O fingerprints did not sample enough isolated O environments. This leads to extrapolation errors when the simulation is moved to 20×20 slabs where oxygen are prone to being in isolated environment (*i.e.,* having no other oxygen atom in its neighbor list). Thus, additional sampling is performed where only one oxygen is added to the surface at different locations. This constitutes the `initial` sub-dataset, which includes 604 configurations.

Note that while this solution suppressed the extrapolation warnings, the periodicity problem is still not fully solved as the Cu atoms still have their own periodic images included in the neighborlist. In the present approach, the movement of the adsorbates are correlated between the periodic images. This may introduce artificial periodicity and limit the degrees of freedom sampled. For our system, such problems were not observed when moving to large, production size slabs. However, in systems where long-range lateral interactions dominates dynamics, the ideal solution, arguably more expensive, is to include slabs with larger lateral dimension, as opposed to limiting the coverage. Fortunately, as we have shown here, introducing a small number of corrective structures can alleviate the problem, so the stronger and more complex short-range interactions can still be sampled with cheaper calculations.

5.3.5 Other problems

Extrapolation could also occur along the vertical direction. While 5 layers is thick enough to converge DFT calculated properties, it is not thick enough for NNP in the sense that none of the Cu atoms are completely surrounded by other Cu atoms in its neighbor list. A sphere of $r_c = 12\,\text{Bohr}$ always includes some vacuum, even on the middle layer. Therefore, when the network potential is evaluated on a thick slab, the Cu-Cu fingerprints extrapolate. An analogous problem exists when the thick oxide is produced, because the 5 layer Cu slab cannot produce a thick enough oxide to create a "bulk oxide-like" environment. It is readily fixed by including a small number of 8 layer slabs in the dataset. These slabs are completely oxidized in MD to generate training set structures that covers the "bulk oxide-like" environments. This constitutes the `thick` sub-dataset, which includes 256 configurations.

Bulk Cu, CuO, and Cu_2O equilibrium MD simulations were added to the training set to ensure bulk properties are well described. High-temperature (1000 K and 2000 K) are used to make sampling more efficient. These constitute the `bulk` sub-dataset, which includes 300 configurations.

5.4 Validation

The average adsorption energies on representative configurations from 3 different phases of oxidation are shown in figure 5.8. Two types of slab structures are used: the $(6 \times 6 \times 3)$ slab and the $(3 \times 3 \times 10)$ slab. These structures are chosen as surrogates for the (20×20) production size slabs, since the latter is prohibitively expensive to calculate directly. Despite the adsorption energy being a stricter test than the error per atom, the final neural network performed well to give adsorption energies within 0.1 eV for most of the configurations. In the worst case, the adsorption energy is within 0.2 eV of the DFT value. Note that this accuracy is close to the inherent error in DFT itself.

Following many machine-learning potential studies in the literature, the ground state potential energy surface is used to parametrize our model. As a word of caution, this

PES is only valid under the Born-Oppenheimer approximation where the motion of the nuclei is assumed to be much slower than the electrons, so that the electrons always have sufficient time to relax to the ground state specified by the nuclei positions. Given our target application of plasma-surface interactions, this approximation may not hold when the ions are moving at velocities far exceeding thermal velocities. The error introduced by using the ground state PES is quantified via the local density friction approximation, which is exact in the limit of an ion moving in a homogeneous electron gas. The resulting friction force are evaluated (see section 5.5.2) to be small compared with the errors of the neural network potential, and hence ignored. Also note that there is no concern about ion kinetic energy exciting the vibrational modes since O atoms are used in the production simulations.

As validation specific to the target plasma oxidation process, the potential energy surfaces that would be explored during oxygen impact on the metal surface are probed by manually placing adsorbates at short distances above the pristine Cu surface. Three surface sites (top, bridge, and hollow) are probed with atomic oxygen(figure 5.9a), vertical (figure 5.9b) and flat (figure 5.9c) O_2 molecules. The structures are illustrated in figure 5.6. Very good agreement between the DFT value and the NNP prediction is found up to 20 eV above the minimum energy adsorption height. Based on these observations, it is safe to conclude that the neural network potential can reproduce Cu-O interaction for the purpose of simulating plasma-metal surface interaction.

As further validation, the potential energy surface of O_2 dissociation on Cu (100) surface were calculated. On the dissociation pathway calculated on the top, bridge, and hollow sites using nudged elastic band method using the reference DFT method, the energies were re-evaluated using the trained NNP potential. On all the NEB images, the energies difference is less than 0.02 eV.

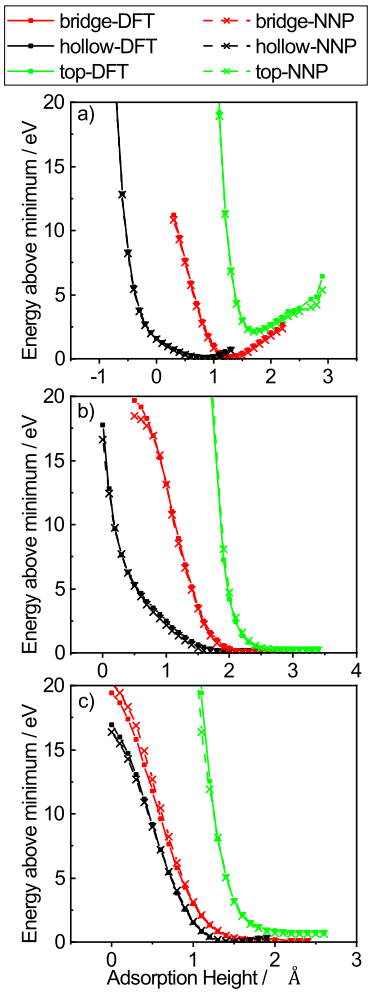


Figure 5.9: Comparison of potential energy surfaces on short-distance, high-energy structures likely present in impact events. **(a)**: a single O atom; **(b)**: a vertical O_2 molecule; **(c)**: horizontal O_2 molecule. The energies are plotted against the distance above the adsorption sites on a Cu (100) surface.

5.5 Discussion

5.5.1 The "sample small, produce large" methodology

One of the greatest promises of machine-learning interatomic potentials is to perform large scale simulations, with PES sampled using (less than 100 atoms) DFT calculations that can be performed efficiently. This approach has a few caveats:

First, the errors can accumulate when going to a large system. The machine learning potentials rely on partitioning the total energy to local, atomic contributions. Each of these contributions E_i are predicted with an variance σ_i, assumed equal to the model error. When using the model on large systems, this error can accumulate. In the case where the errors are completely uncorrelated, a system with N atoms have an variance of $\sqrt{N}\sigma_i$. For a large system this can be significant. Therefore, the energy predictions from large, uncorrelated structures may not be comparable. In DFT, error cancellation is heavily relied upon for reasonable energy differences. Here, it is less applicable due to the random nature of the errors.

On the other hand, if only local dynamics is the target, as is the case in chapter 6, the error in the forces do not scale with system sizes and the dynamics are correctly captured as long as the model reproduces the PES well, independent of system sizes. This brings naturally the question of ensuring model transferability across sizes, *i.e.* how do we know the model gives the same per-atom error on a large, production stucture as on the training structure?

One obvious approach is to generate slightly larger structure that can still be treated in standard DFT packages. This is the approach adopted here when the validation is performed on slabs of different shapes in figure 5.8. This can also be done on production-size slabs using efficient scalable DFT packages. This is the ultimate test, but can only be performed on a few selected structures.

Another possibility relies on comparing atomic environments encountered in production runs to those already seen in the dataset. The simulation can be stopped if the model is

operating outside the range explored in the training set. This is reminiscent of active learning approaches where, in addition to the warning, the offenders are sent to DFT and the results added to dataset.

At the core of such methods is an extrapolation measure. There are 2 ways of doing this. The first, the so-called query-by-committee method, requires training a "committee" of models from different random starting points. On the test structure, each member model performs a prediction on the test structure and the variance among the prediction is taken as the extrapolation measure. This approach is conceptually simple and can be applied to any machine learning framework, but retraining the committee frequently may be a practical bottleneck. The second approach relies on the atomic environment featurizer themselves. For the kernel methods, the prediction variance is defined when the kernel itself is defined. Therefore, kernel methods (*e.g.* GAP) has natural built-in bayesian extrapolation measure. The downside is that inferrence time scales with the size of the training dataset. For feature vector-based methods, the extrapolation grade based on the MaxVol algorithm can be used. However, this algorithm rely on the volume spanned by active set, therefore suffering from being unstable with respect to outliers in the training set. A single outlier datapoint can greatly expand the said volume, so the algorithm considers points near it to be "intrapolating", even when such regions of the PES are not sampled adequately. In the present contribution, a similar, but more simplistic approach is used when the range of the feature vectors is used directly, suffering from similar drawbacks.

5.5.2 Error introduced by the adiabatic assumption

As discussed in the main text, the neural network potential is fitted to the ground state potential energy surface. Dynamics on this potential are valid as long as the nuclear and electronic degrees of freedom is separated (*i.e.* the adiabatic assumption is valid). However, situations exist where the nuclear motion can lead to electronic excitations, and vice versa. On metals, due to the absence of band gaps, it is known that electron-hole pairs can be created by the motion of adsorbates, leading to what is known as nonadiabatic adsorption.

To quantify the error introduced by neglecting this energy dissipation mechanism, we have calculated the forces felt by an oxygen ion moving above an Cu (100) surface in the local density friction approximation (LDFA).[77] In LDFA, the friction force on ion is directly proportional to the velocity v, with a density dependent scalar friction coefficient η:

$$\vec{F} = -\eta\vec{v} \tag{5.1}$$

The steps taken to evaluate the friction force is outlined below. First, the densities are converted to the Wigner-Seitz radius as:

$$r_s = (\frac{3}{4\pi\rho})^{1/3} \tag{5.2}$$

Then, the correspondence between the Wigner-Seitz radius and the friction coefficient and an oxygen atom moving in a homogeneous electron gas is obtained from Ref.[156]. The tabulated results are interpolated via a smoothing spline to obtain a continuous relation, as shown in figure 5.10. The full-electron density traversing vertically normal along a Cu (100) slab through one of the surface atoms is obtained via the PAW method and shown in figure 5.11. The sampling line traversed through the atomic cores, and the plotted density is truncated to show only the relevant regions. Visual aids were inserted in figure 5.11 to indicate regions within 1 Å of the nuclei. In MD impact simulations, the Cu-O distances were never observed to be shorter than this cutoff distance, so the forces inside this shaded region is irrelevant. The friction coefficients are shown in the left axis, and the friction force on an oxygen ion with 20 eV kinetic energy is shown on the right axis. It can be seen that the friction force never exceeds 0.08 eV/Å, which is less than the error in our neural network potential. Hence, we conclude that friction due to electronic friction is insignificant for our purposes.

5.6 Conclusions

A machine learning potential is developed and validated for the Cu/O system that covers atomic and molecular adsorption with initial kinetic energies up to 20 eV. While the potential was used to study plasma oxidation of copper, the approach used is quite general and could

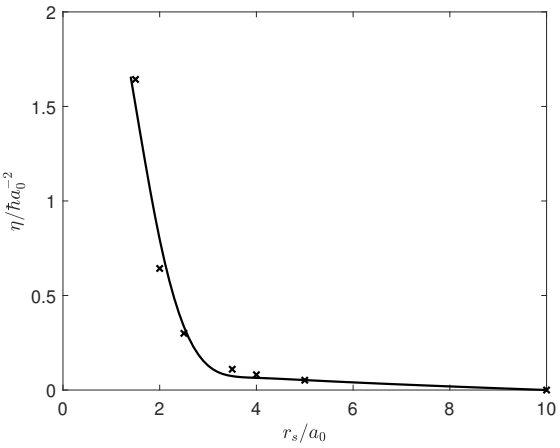


Figure 5.10: Tabulated correspondence between local electron density (expressed as Wigner-Seitz radius) and the friction coefficient. The curve corresponds to spline interpolation. The friction coefficient for $r_s > 10$ is assumed to be 0.

be extended to other binary metal-modifier interactions, limited only by the ability of the underlying electronic structure method to produce a large set of consistent and accurate training data.

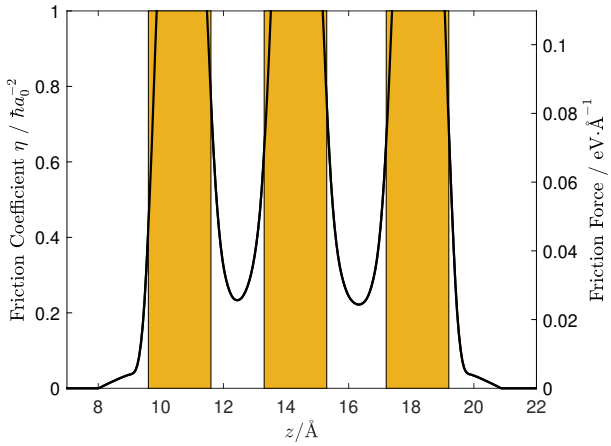


Figure 5.11: Calculated LDFA friction coefficients and forces for oxygen atom moving through an Cu (100) slab with a kinetic energy of 20 eV, plotted for density values along a vertical line normal to the surface and passing through the Cu nuclei. Note: this overestimates the actual density experienced by the oxygen ion, hence the actual forces are lower than what is suggested here.

CHAPTER 6

Molecular Dynamics Simulation of Plasma Oxidation of Copper

6.1 Introduction

In chapter 5 the development of the general Cu/O interaction potential discussed. In this chapter, the application of this potenital to plasma oxidation of copper to predict the influence of processing conditions such as substrate temperature, plasma power, and chamber pressure. Suggestions on improving the existing experimental process are given on the basis of the influences found.

After replacing aluminum decades ago, copper became the universal material of choice for interconnects in integrated circuits.[39] Traditionally, copper was introduced into the back-end-of-line (BEOL) processing by a combination of the Damascene process which plates copper onto patterned interlayer dielectrics, followed by a chemical mechanical polishing (CMP) process.[51] However, continuous scaling down in the push to follow Moore's law is beginning to render this harsh process unsuitable at the lowest levels of metallic interconnects, due to requirements of high-aspect-ratio filling uniformity and cracking induced by CMP.[179]

In response to these problems, the semiconductor industry is looking for an atomic-level engineering approach that is simultaneously capable of achieving high selectivity and high directionality. One of the most promising techniques for this purpose is atomic layer etching (ALE).[105] To etch metallic copper, a specific variant, the plasma-thermal ALE process can be used. It is composed of two steps (figure 5.1): 1) controlled conversion of the surface layers of copper to copper oxide under oxygen plasma and 2) reaction of the oxide layer

with an etchant molecule (*e.g.* HCOOH) to form volatile complexes, etching the oxide layer and stopping at the metallic part.[173] While plasma is required to impart the inherent directionality of the ions accelerated by the electric field to the device profiles, the process leads to relatively thick activated layers without being self-limiting, preventing a finer grain of control over etch per cycle (EPC) from being achieved.[126]

Copper oxide is one of the enabler materials of modern technology. Stable under ambient conditions with an optimal band gap of $\sim 2.0\,\mathrm{eV}$, it is a highly desirable and readily available material for photocatalysis and photovoltaic cells. Incidentally, the oxidation of copper is one of the most heavily studied oxidation processes. The Cabrera-Mott theory of metal oxidation was conceived partially as an explanation for the formation of thin copper oxide films. While the classical theory was shown to be inaccurate as understanding improved,,[116] copper oxidation is now considered somewhat of a model system to understand formation of thin metal oxides in general.[106]

The oxidation of low-index copper surfaces has been extensively studied under surface science conditions.[95,112,124,153,170] Oxygen is known to adsorb dissociatively on the (100), (111), and (110) terminations, forming well-characterized configurations and reconstructions at temperatures below room temperature. As coverage increases, the oxide grows as nanoislands on all three terminations,[45] as opposed to the uniform layer growth assumed in earlier works. The size and shape of the islands depend on the orientation of the surface, the conditions, and the oxygen dose. In conditions departing from the ultra high vacuum towards the more realistic pressures, temperatures, and time scales, the process is far less well understood. There is little consensus on the form of the rate law at temperatures near or above room temperature, and the thickness and the composition of the native oxide is not well understood yet, nor the influence of temperature, moisture, *etc.*[135]

In addition to the difficulties that make understanding thermal oxidation challenging, the interaction of metal surfaces with the plasma is unique in having energetic particles impacting the surface.[133] The particles used in the corresponding experimental process (the "target process") have kinetic energies in the order of $10\,\mathrm{eV}$,[49,149] high above those possible from thermal fluctuations, yet too low to cause ion implantation or sputter the substrate atoms.

Unlike the case for higher energies (>1 keV) where simple screened Coulomb potentials coupled to binary collision approximation give very reasonable results, here an accurate description of the potential energy surface is still relevant, since the kinetic energy is within an order of magnitude of the strength of a typical chemical bond. In other words, the plasma serves to accelerate the underlying chemistry rather than obscuring it.

6.2 Methods

To successfully model the large system with a heavy machine-learning potential, a special setup is created that adaptively includes more atoms into pair interaction calculation and time integration only when necessary. Although the simulation has $(20 \times 20 \times 50 = 2 \times 10^4$ atoms at the beginning, only the top region (around 5 layers) is integrated. The interaction among the Cu atoms deep inside the bulk lattice is not calculated, and their positions are not updated by the time integration. The input script periodically checks for the coordinate of the deepest-penetrating oxygen atom, and includes more copper atoms into the integration and interaction calculations, so as to maintain a constant reservoir of pristine copper atoms ahead of the progress of the oxidation front. In the current implementation, this check is done every 10 impact events, and is found to be more than adequate. Using this setup allows the simulation to have effectively much smaller number of atoms and thus proceed much faster, without consuming the entire copper slab. It is noted that this method does not corrupt the trajectory because ion impact is known not to cause Cu lattice defects *a priori*.

A similar setup is created to recede the height at which new ions and neutrals are deposited. The goal is to minimize the travel distance of newly added oxygen species before they reach the surface. All 3 regions depicted in figure 6.1 are receded adaptively. In this case, the recede is based on the z-coordinate of the topmost copper atom.

To accurately capture impact event, the integration time step must be very small to account for the very fast ion velocity. However, when the impact is finished, using such small time steps are very inefficient. Therefore, the time step is made to scale with the

maximum atomic velocity present in the system. The time step is rescaled periodically (every 10 steps) such that no atoms move more than a distance of 0.01 Å in each time step. This distance is the maximum distance at which total energy is still conserved upon ion impact, established by separate testing.

To preserve the thermal spikes created by the plasma impact, only the regions near the edges are thermostated, as shown in figure 6.1 a). Hence, most of the atoms under the impact approximate the microcanonical ensemble, while the system as a whole approximate the canonical ensemble.

In the actual process, the pressure in the chamber is maintained by a line of vacuum pumps. In the present simulation, pressure is not considered. However, a mechanism to maintain the number of oxygen molecules in the system necessary to prevent the surface from being fully covered due to recombined oxygen molecules (O_2). After every ion impact, the simulation is paused and the atomic configuration extracted. This configuration is analyzed to obtain the indices of the recombined O_2 molecules. A certain fraction of such bonded pairs are removed, and the simulation is continued. This on-the-fly analysis is enabled through the Python interface of LAMMPS. Likewise, in the simulations involving 400 eV ions, the sputtered Cu atoms are removed using the same approach. This is necessary because the training data did not include isolated Cu atoms and metal re-deposition is beyond the scope of this work.

6.3 Results and Discussion

With the trained and validated potential, large-scale molecular dynamics simulations are performed. Figure 6.1 shows the setup used to perform long-time simulations. Cu(100) slabs are constructed from their DFT-optimized bulk unit cell. The vertical space in the simulation box is divided into desorption, addition, buffer, and the substrate itself, as shown in figure 6.1b). The oxygen plasma is modeled as a mixture of atomic ions and neutral radicals, with ion to neutral ratios actively controlled as a parameter. The O_2 molecules, although correctly described by the potential, are not included because reactivity of the

plasma is dominated by the ions and neutral radicals. Atomic oxygen ions are added to the cell every 10,000 steps, roughly corresponding to 5 ps, with a uniform z-velocity calculated from their specified kinetic energies E_k. The x, y-velocities are specified randomly from a Gaussian distribution normalized to the root mean square velocity ($v_{\mathrm{rms}}(T_s)$) at the specified temperature T_s. Slow neutrals are added at an interval set according to the specified ratio (n/i) relative to the ion addition frequency. For example, a ratio of $n/i = 10$ would indicate that 10 neutrals are added per ion. The three velocities components are initialized similarly to the x, y-velocities of the ions. Such ratios are typical given the experimental conditions.[173] Apart from the velocity, the charge is not explicitly modeled since most of the ions are neutralized by Auger electrons before reaching the surface.[133] All the ions are deposited essentially vertically downward: a collision-free presheath is assumed. Frequently, oxygen atoms recombine on the surface and desorb as O_2 molecules. These molecules are removed as soon as they reach the desorption region. No copper atom has been sputtered away from the substrate during the MD simulations.

These slabs have square unit cells with 5 nm lengths in each lateral direction, and 12 nm thickness to ensure that the oxidation never reaches the bottom layers. The horizontal direction is divided into the center region where an NVE ensemble is used, and the edge regions where an NVT ensemble is used. This maintains the system at the thermostat temperature but avoids corrupting the thermal agitations created by the ions. In other words, the kinetic energy imparted by ions has to propagate via phonons to the thermostated region, where it is dissipated via the thermostat.

There are three variable parameters in the simulation of direct experimental significance: substrate thermostat temperature (T_s), ion kinetic energy (E_k), and the ratio of neutral to ions (γ_{ni}). The temperature corresponds well to the temperature of the wafer. The ion kinetic energy is affected simultaneously by the plasma power and the applied bias. The ratio of neutral atoms to ions is affected by the chamber pressure and the plasma power. Although the quantitative relations between simulation parameters and experimental parameters are not known without careful plasma diagnostics, the trend is clear: Higher plasma power and higher pressure correspond to more ions of higher energy, at the possible expense of decreased

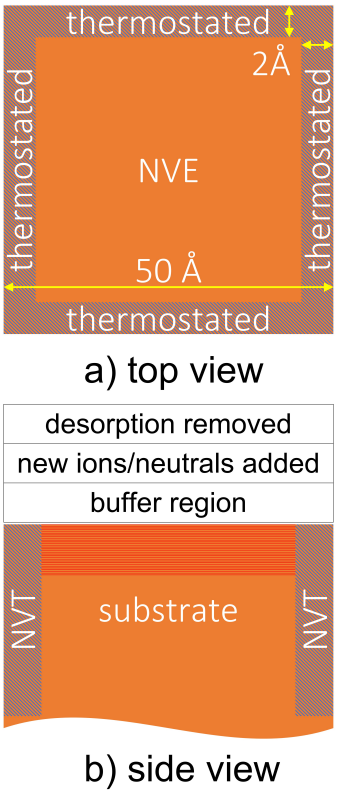


Figure 6.1: Overall setup of MD simulations. The production simulations used (20×20×50) slabs. The schematic here is not drawn to scale. **(a)**: the lateral cross section is divided into thermostated edges and non-thermostated centers; **(b)**: the vertical space in the simulation box is divided into the desorption removal, new atoms addition, buffer, and the substrate regions

directionality.

In figure 6.2, the evolution of the Cu(100) slab structures under $T_s = 293\,\mathrm{K}$, $E_k = 10\,\mathrm{eV}$, $\gamma_{ni} = 10$ is shown. Changing these parameters does not affect the general stages of oxide growth discussed here. Starting with a pristine surface, the oxygen atoms adsorb to form an overlayer. Higher energy ions transfer their energy to the surface atoms but do not penetrate into the sub-surface. At this stage, the sticking coefficient S is very high, and almost all the atoms added adsorb to the surface. As shown, the adsorption is not uniform on the surface. The adsorbate atoms tend to migrate between different sites initially, during which they impinge upon the copper atoms on the top layer and transfer the kinetic energy to the surface. Once their kinetic energies fall below the diffusion barrier on the (100) surface, they settle down at a four-fold hollow site. As the O coverage is increased, the bonding between the first layer of copper atoms and the layer underneath is weakened. Thus, a "peel off" phenomena is observed where small flakes of the copper oxide layer would detach slightly from the layer underneath. This detachment lowers the barrier for oxygen diffusion to the subsurface sites. In the process, a corrugated and disordered surface is created, with islands of oxides and basins of pure copper (figure 6.2, $t = 56\,\mathrm{ps}$). Corrugation is significant in oxide growth for two reasons. First, it allows the oxygen to migrate into the copper lattice before the surface adsorption sites are completely occupied. In figure 6.2, $t = 170\,\mathrm{ps}$, when the oxide is around 3 atomic layers thick, there remains a region on the surface not completely covered by oxygen. Second, the same process is repeated at the interface between the oxide and the copper underneath: Cu-O binding weakens the binding of Cu atoms to the copper lattice, lifts the Cu atoms out , allowing oxygen atoms to diffuse through to the vacancy created.

Beyond a few atomic layers, the growth of oxide slows down rapidly and transitions to a diffusion- and/or interface-reaction-limited process, characterized by an amorphous copper oxide layer resistant to further oxygen adsorption. In figure 6.2, the oxide growth between 1587 ps and 4892 ps only amounts to approximately two atomic layers, in stark contrast to the faster growth earlier. The rate of oxygen incorporation decreases, and the rate of oxygen recombination and molecular desorption increases. Figure 6.3a) shows the growth of oxide

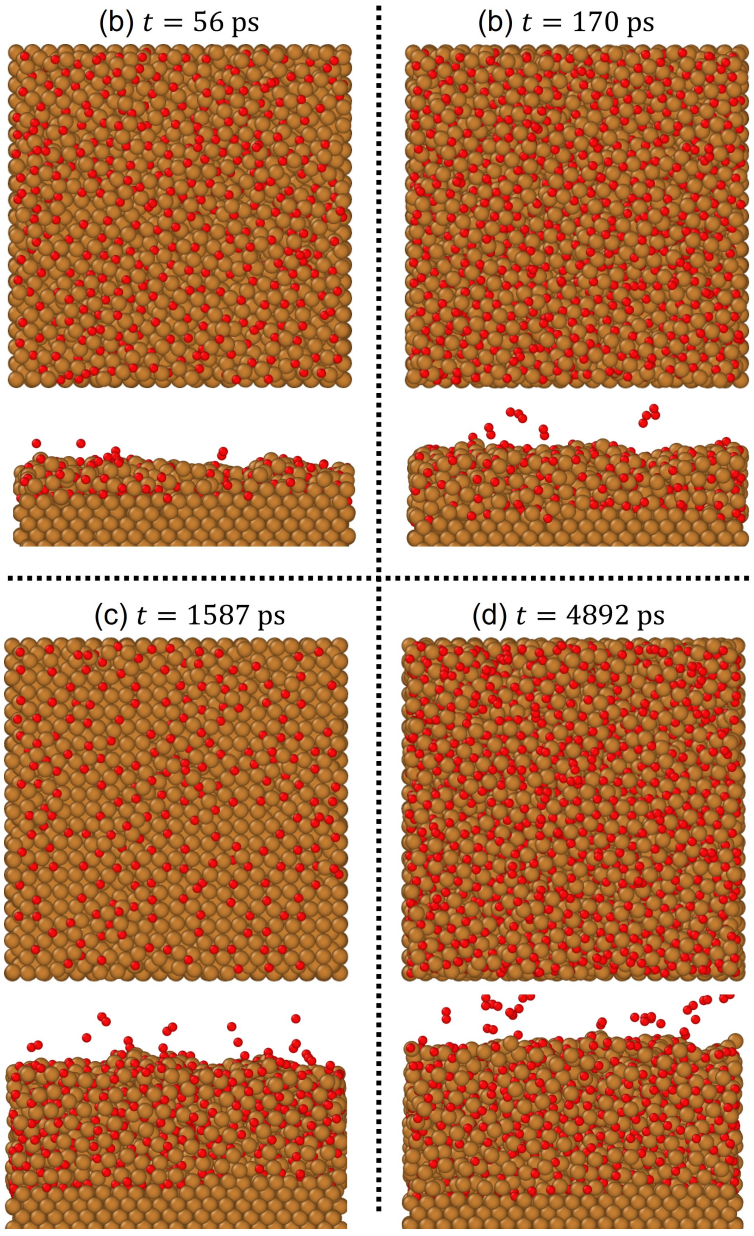


Figure 6.2: Snapshots along the oxidation trajectory of Cu(100) slab at four different points in time. **(a), (b)**: the very early stage where the growth is determined by the availability of oxygen (controlled by the neutral to ion ratio since ions are deposited at fixed step intervals). Between **(a)** and **(b)** the adsorption sites become exhausted, and the growth transitions to much slower the kinetic energy limited regime. Between **(c)** and **(d)**, the oxide thickness grew only by approximately 2 atomic layers, despite a time elapse of 3.3 ns.

thickness with time, and figure 6.3b) shows the amount of oxygen incorporated into the slab. The details of determining this thickness from the MD trajectory are given in appendix C.1.1. The recombined O_2 molecules can be seen in the side views shown in figure 6.2. These molecules can remain physisorbed for picoseconds, before they eventually acquire enough energy from thermal fluctuations. Note that the simulation is likely to overestimate the number of these molecules because the time between impacts is much shorter than that in reality. In reality, these O_2 molecules would have had much more time to desorb.

Comparing the results for the same temperature, at least 2 regimes of growth can be identified: the fast regime at first, and the slow regime later. Two strands can be separated from the fast regime, each corresponding to different neutral-to-ion ratios, and independent of the kinetic energy of the ions. As exposed adsorption sites become occupied, growth slows down significantly. No limiting thickness is observed. Instead, the oxides continue to grow slowly at a rate of around 0.15 nm/ns. The growth rate in this regime seems to be determined by E_k, where the thickness profiles gradually separate into two strands distinguished by a faster growth at the larger E_k of 20 eV versus the slower growth at 10 eV.

The effect of the temperature of the substrate thermostat T_s can also be seen by comparing curves with the same line color and symbol, but with different symbol colors. Higher substrate temperature is seen to promote oxide growth by postponing the transition to the slower regime. Its effect is similar to but less pronounced than that of the ion kinetic energy.

The two regimes can be rationalized by considering the physical processes at work. Initially, when the pristine surface is exposed to oxygen plasma, the high-energy ions have a large free energy driving force to adsorb. Hence, it is expected that most of the impinging ions adsorb. Moreover, the ions do not have enough kinetic energy to penetrate even below the first layer. In figure 6.2, from the first snapshot, it can be observed that all oxygen atoms are stopped at the top layer. Therefore, the kinetic energy of the ions are transferred to the Cu lattice and dissipated through lattice vibrations into the bulk copper. In the simulation, they dissipate to the thermostated regions, without creating lattice or surface defects in the process.

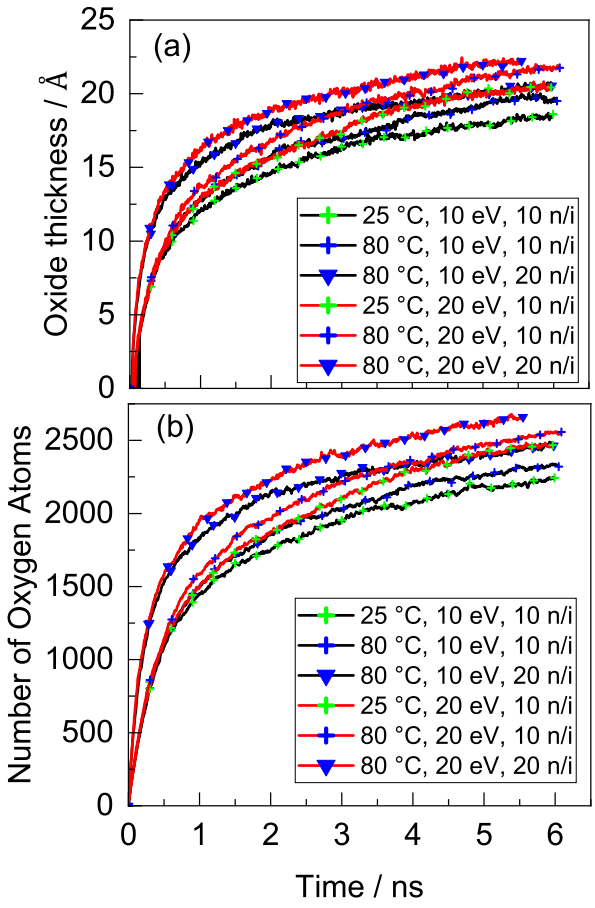


Figure 6.3: Thickness and oxygen incorporation as functions of simulation time for all six conditions explored. The growth curves start as two distinct strands, where the grouping is determined by the neutral to ion ratio (indicated by the shape of the marker, cross vs triangle). The strands quickly splits apart based on the kinetic energy of the ions. The curves corresponding to $E_k = 20\,\text{eV}$ (red) grows faster than those corresponding to the $E_k = 10\,\text{eV}$ (black), eventually separating into another pair of strands where grouping is determined by E_k. The trend of oxygen incorporation is very similar to the thickness.

As the oxide grows, fewer and fewer surfaces Cu atoms with low coordination are exposed, reducing the sticking coefficient. Oxide growth becomes increasingly limited by the migration of O atoms into the Cu lattice and the reaction of oxidation at the oxide-metal interface. Both reactions are limited by the energy available to the oxygen atoms at the interface to overcome the energy barrier. In the thermal oxidation process, the only source of energy would be the random thermal fluctuations. Hence, growth rate is expected to depend mainly on the substrate (thermostat) temperature. In a plasma-enhanced process, the impinging ions act as additional source of thermal energy. The thermal spikes created by the ions propagate through the oxide layer and the Cu lattice, increasing the kinetic energies of the atoms in its path. Only a fraction of the energy is transferred to the next atom in each collision event. The thermal spike heats the atoms along the path of its cascade to kinetic energies of several orders of magnitude higher than the Maxwell-Boltzmann distributed thermal fluctuations would allow. Such high energies would overcome reaction barriers easily. However, it is demonstrated next that such highly concentrated energy is only available to atoms within a few layers of the impact event, because of the fast dissipation. This is the case with the amorphous copper oxide layers, where non-symmetric environment means that neighboring atoms do not always exist in the direction of propagation, hence the energy is more likely to dissipate.

To further explore this point, impact simulations were performed on an identical oxide surface with a single ion. In these simulations, a snapshot is taken towards the end of the MD trajectory shown in figure 6.2 as the starting point, corresponding to an thick ($> 2\,\mathrm{nm}$ oxide above the Cu substrate. 500 replica simulations were run in which one oxygen ion is launched at the surface with a randomly chosen lateral starting coordinate. The trajectory is followed for $\sim 3\,\mathrm{ps}$. In figure 6.4a-b, an attempt is made to calculate the maximum depth of the atoms agitated by the heat spikes and plot their distribution as a histogram. The penetration depth is defined as the depth of the deepest atom that has a kinetic energy with probability less than $f_{\mathrm{cut}} = 1 \times 10^{-5}$ in the Maxwell-Boltzmann distribution and is connected by a series of collision events to the impact site. The specific cutoff probability is motivated by practical purposes: using a higher probability leads to random thermal fluctuations being mistakenly

counted as ion-induced agitation (see appendix C.1.3). Hence, the quantitative value of the penetration distance is less significant than the trend, which clearly shows that increasing the kinetic energy from 10 eV (figure 6.4a) to 20 eV (figure 6.4b) shifts the distribution to larger distances. This observation is reminiscent of the ion implantation technique, where both the projected range and the straggle increase as the kinetic energy increases. Thus, the ions serve as a transient, anisotropic, high intensity energy source to help thermally activate the oxidation reaction. This contrasts with the thermal energy, which is persistent, isotropic, and low in intensity.

In figure 6.4c, a more direct test is performed where either ions or neutrals deposition were stopped mid-way in the simulation. Comparing the growth trends shows that oxide film growth slows significantly when either ions or neutrals are stopped. In particular, stopping the ions slows growth immediately, whereas the effect of stopping the neutrals does not become apparent until ~ 2 ns later. This is fully consistent with our observation that the growth is accelerated by the kinetic energy imparted by the ion. Because the time scale of thermal energy dissipation is short (lasting less than 1 impact event), the directional, intense energy is depleted quickly, which is immediately reflected in the growth trends.

The effect of ion energy becomes even clearer when the ion energy range is increased 20-fold to 400 eV. This is enabled by the physics-informed NNP+ZBL model discussed in section 5.2.2. The growth curve obtained with $T_s = 353\,\mathrm{K}, n/i = 100$, with ion energies set to 0 eV, 100 eV, 400 eV are shown in figure 6.5

On the basis of the distinct nature of heat sources, a third regime is likely to be present where the growth rate is determined by the substrate temperature. However, extending the simulation long enough to discern this regime is prohibitively expensive. If the rationale holds, this temperature-limited regime would be present when the oxide grows beyond the thickness affected by the thermal spikes. As a warning, note that the 6 ns of simulation time covered does not correspond to the experimental time, since the deposition rate is much higher than what is likely under experimental pressures (see the comparison in appendix C.2).

Voronoi tessellation averaged bond orientational order (BOO) parameters is used to study

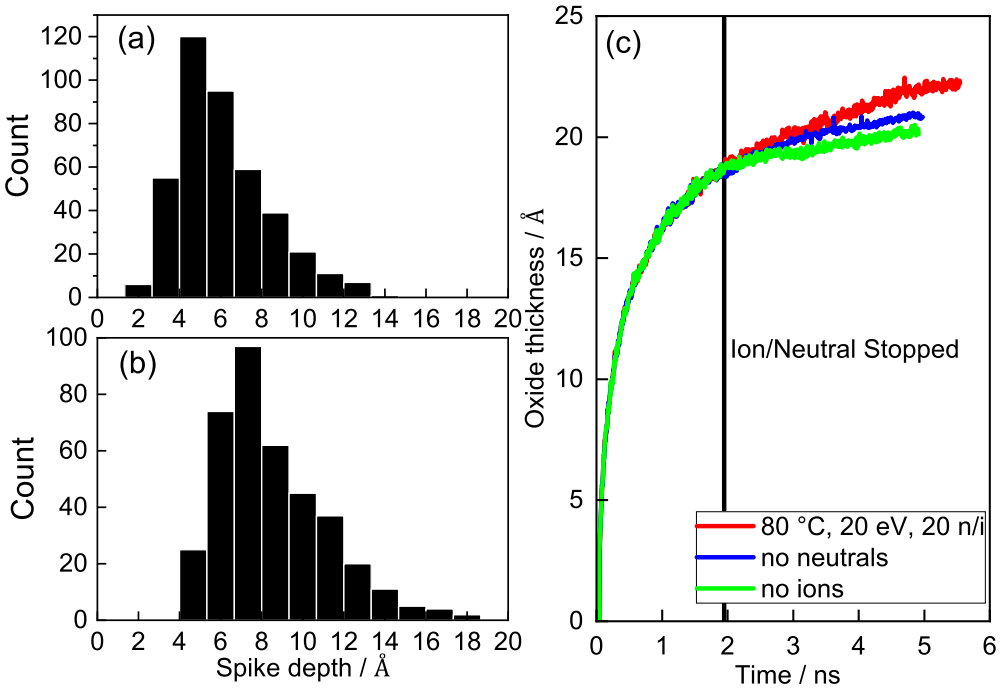


Figure 6.4: Distribution of the projected range of thermal agitation on copper oxide slabs. The oxide slab is obtained by first taking an thick oxide slab resulting from the long-time simulation. Oxygen ions are added to the slab with different random seeds determining the lateral position. 500 replicas are performed for each kinetic energy levels. **(a), (b)**: results from using 10 eV and 20 eV atomic oxygen ions respectively. **(c)**: growth curve after stopping ion or neutral deposition at 1.91 ns

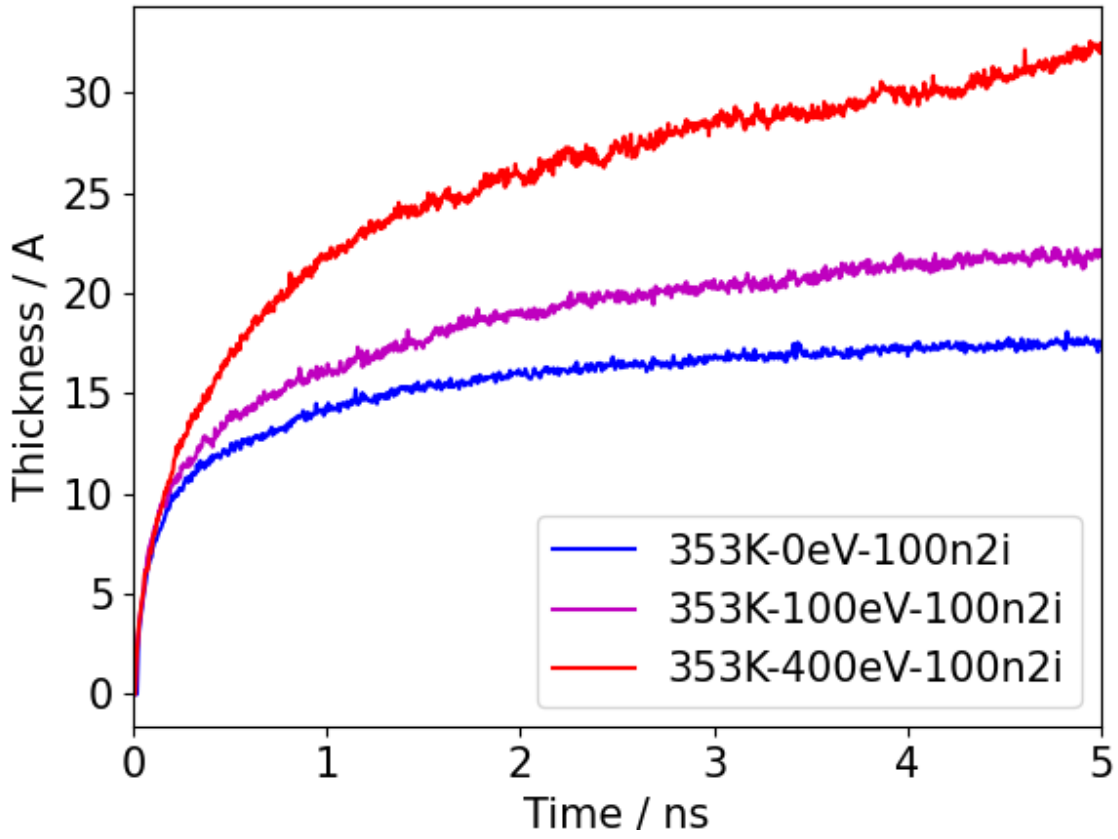


Figure 6.5: MD growth curve obtained at $T_s = 353\,\mathrm{K}$, neutral-to-ion ratio = 100, and ion energies of 0, 100, 400 eV.

the crystal structure of the copper oxide layers.[12,97] It has become a commonly used tool used to identify different crystalline phases and clusters. Based on spherical harmonics, these encode the orientations of the bonds around each atom, as given in appendix C.1.2. Shown in figure 6.6a are the resulting $\bar{q}_4$ and $\bar{q}_6$ parameters for high-temperature equilibrium MD of supercells of the two copper oxide bulk crystals Cu_2O and CuO. The two structures are clearly distinguishable on this figure. CuO has two clusters with similar $\bar{q}_6$ parameter values around ~ 0.40, but different $\bar{q}_4$ values at ~ 0.35 and ~ 0.70, corresponding to O and Cu atoms, respectively. On the other hand, Cu_2O has two clusters with similar $\bar{q}_4$ values at ~ 0.50, but different $\bar{q}_6$ values centered at ~ 0.5 and ~ 0.3, respectively, for O and Cu. The spread is due to thermal fluctuations as a result of the high-temperature equilibrium MD.

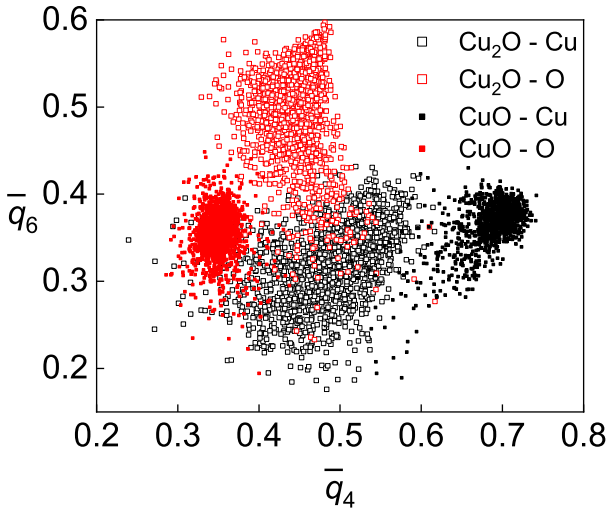


Figure 6.6: Bond orientational order parameters for bulk CuO and Cu_2O, used as references for distinguishing local environments of the oxide slab obtained from MD. The distinguishing feature is similar $\bar{q}_4$ parameters and different $\bar{q}_6$ for CuO, but the opposite for Cu_2O.

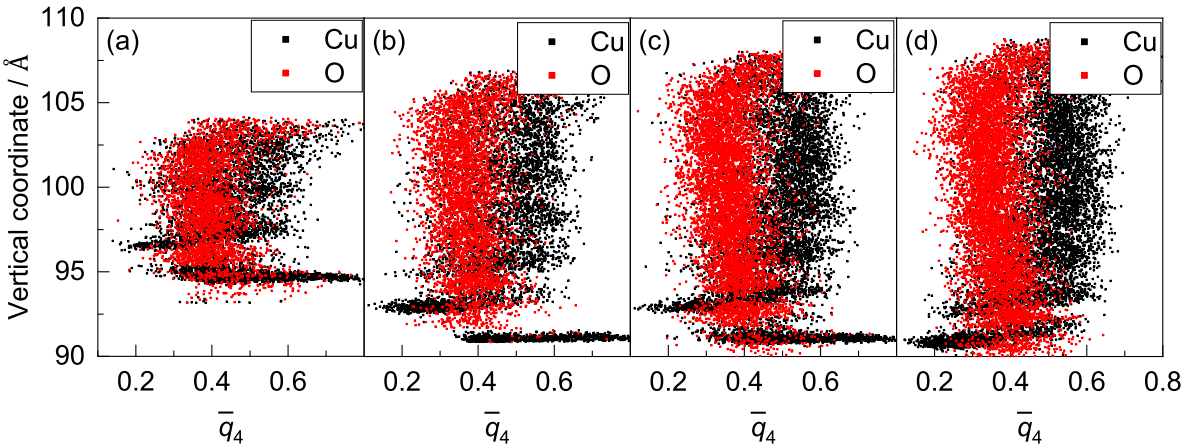


Figure 6.7: Bond orientational order parameter $\bar{q}_4$ for a range of oxide thicknesses. The structures used are taken from the $T_s = 353\,\mathrm{K}, E_k = 20\,\mathrm{eV}, n/i = 10$. The z-coordinates of atoms are plotted against the $\bar{q}_4$ values, showing a gradual horizontal separation of the Cu atoms from O atoms that becomes more pronounced as the oxide increase in thickness. The $\bar{q}_6$ parameters are relegated to supplementary information figure C.1 since the clusters overlap one another. The sharp Cu clusters toward the bottom of each panel are due to metallic copper atoms near the oxide-metal interface.

In figure 6.7, the calculated BOO parameter are plotted against their vertical coordinate (z-coordinate). For clarity, only $\bar{q}_4$ is shown. The $\bar{q}_6$ parameters of Cu and O overlap each other and are shown in figure C.1. The graphs used in figure 6.7 are extracted from $T_s = 353\,\mathrm{K}, E_k = 10\,\mathrm{eV}, n/i = 10$ simulations and re-equilibrated at 353 K to remove thermal spikes that may have resulted from recent impact events. The structures clearly illustrate the growth of an oxide-like local environment, distinct from the layered pristine copper atoms toward the bottom of each panel. Furthermore, the parameters indicate the dominance of local CuO-like environment, as indicated by the segregation of the $\bar{q}_4$ parameters derived from Cu and O into clusters at values of 0.3 and 0.6, respectively. Such segregation is not observed in the very thin oxide of figure 6.7 (a), but becomes pronounced only as the thickness increases. The position of the clusters stabilize beyond a thickness of 2 nm. Comparison with figure 6.6 suggests that the resulting oxide after long exposure to oxygen plasma is of the structure and composition of CuO. This result is consistent with experimental X-ray photoelectron spectroscopy study,[173] although in that case the Cu_2O is not ruled out. The $\bar{q}_6$ result suggests no Cu_2O structure is present (figure C.1). The overall stoichiometry of the oxide film is close to 1 (see figure 6.14 for local atomic concentrations), which agrees with the analysis of the local environment.

In figure 6.7, the point clouds of Cu and O atoms clearly cluster in the $\bar{q}_4 - z$ plane. Moreover, there are hints of clustering in the z-coordinate as well. It is suspected that the repeated ion impact would cause repeated local heating-and-cooling, similar to annealing. Another set of simulations was performed with the ion addition frequency maintained at 10000 steps per ion and neutral to ion ratio at 10, but on the smaller (10×10) Cu (100) slab. In figure 6.8a, BOO analysis similar to that of figure 6.7 is performed on trajectory obtained using the setup $T_S = 353\,\mathrm{K}, E_k = 10\,\mathrm{eV}, n/i = 10$. The structure selected is at the end of a 1 ns trajectory. Effectively, the oxygen flux has increased by a factor of four, compared with the simulation on the (20×20) slab. In other words, the oxygen fluence at the 1 ns simulation time is comparable to that at 4 ns on the (20×20) slab. The resulting BOO parameters in figure 6.8a show clear clustering in the z-coordinate. The structure itself shows quasicrystalline CuO regions within ~ 5 layers above the oxide-metal

interface, as also captured by the BOO parameter. Above the quasicrystalline region, the film is amorphized by presumably repeated high-energy ion impacts. The segregation in z allows density-based spatial clustering of applications with noise (DBSCAN) clustering to be performed on the clusters. The centers of these clusters are shown in table 6.1, where the $\bar{q}_4$ and $\bar{q}_6$ coordinates of O and Cu clusters match well with the reference coordinates in figure 6.6 for CuO, confirming that the local environment closely resembles CuO. The top layer of this film structure remains amorphous, due to roughening effect of ion impacts.

The z-coordinates of the clusters are labelled in figure 6.8. The distances between successive Cu and O clusters are calculated to be (1.27 ± 0.02) Å, whereas the interlayer separation in bulk CuO is 1.28 Å along the [001] direction. This is another indication that the structure is crystalline. Comparing to the simulation performed with the identical MD $T_s, E_k, n/i$ parameters but lower deposition rate on the (20×20) surface, at equivalent fluxes, the film thickness is reduced. The same thickness analysis performed on the (10×10) slab shown in figure 6.8 gives thickness of 16.7 Å, whereas at the point of equivalent oxygen fluence (4 ns), the (20×20) slab has a thickness of 18.6 Å, as shown in figure 6.3. Considering the quasi-crystalline nature of the film formed on the (10×10) surface, this is reasonable since diffusion across the crystalline region is much more difficult.

It is clear that the (20×20) surfaces, with its lower oxygen flux, correspond much better to experimental conditions. However, at the moment, it is unknown whether further decreasing the flux can yield a different growth trend and film morphology. The results here suggest that 1) the unrealistic ion and neutral fluxes used in plasma-surface simulation can have a significant impact on the morphology of the simulated film and 2) it might be possible to control film crystallinity, in addition to thickness, by controlling the plasma parameters. As technical notes, it is acknowledged that 1) the observed crystallinity could be induced by the increased artificial periodicity introduced by the smaller lateral dimension. This possibility can be ruled out by performing a set of simulations on the (20×20) surface with 4-fold increase in the deposition rate; and 2) the change in the BOO parameters near the top of the oxide slab seen in figures 6.7 and 6.8 does not indicate a distinct structure, but is a manifestation of the unreliable BOO parameters on surfaces due to the tendency

Table 6.1: Centers of clusters in as given by DBSCAN analysis, showing alternating layers of Cu and O atoms.

q_4	q_6	z-coordinate / Å	Element
0.56	0.35	94.12	Cu
0.36	0.37	95.41	O
0.61	0.37	96.70	Cu
0.34	0.35	98.00	O
0.61	0.37	99.29	Cu
0.33	0.34	100.50	O
0.60	0.38	101.80	Cu
0.33	0.34	103.06	O
0.59	0.38	104.34	Cu
0.34	0.33	105.61	O

of the Voronoi tessellation definition of the nearest neighbors to overestimate the number of neighbors.[164] Lastly, the above results demonstrate that the accurate machine learning potential is able to qualitatively capture the phase change from metallic Cu to amorphous CuO_x, and finally to crystalline CuO. Moreover, it did so without any prior assumptions or external bias potential favoring certain phases. This capability has not been demonstrated by classical force fields on this system but is essential to model this complex, polyvalent oxidation process. Whether the transition occurs at accurate energy and time scales remains to be seen by further studies.

An attempt was made to understand the extent to which oxidation can be described by a single activated process. A "per-particle temperature" is calculated from the instantaneous particle kinetic energy according to the equipartition principle and used to obtain the individual contribution $\frac{dL}{dt}$ as an Arrhenius-type relation (equation (6.1)), where L is the thickness of the oxide layer. The contributions of the individual particles are averaged to obtain the growth rate $\bar{r}(L)$, as a function of film thickness (equation (6.2)). Therefore, the

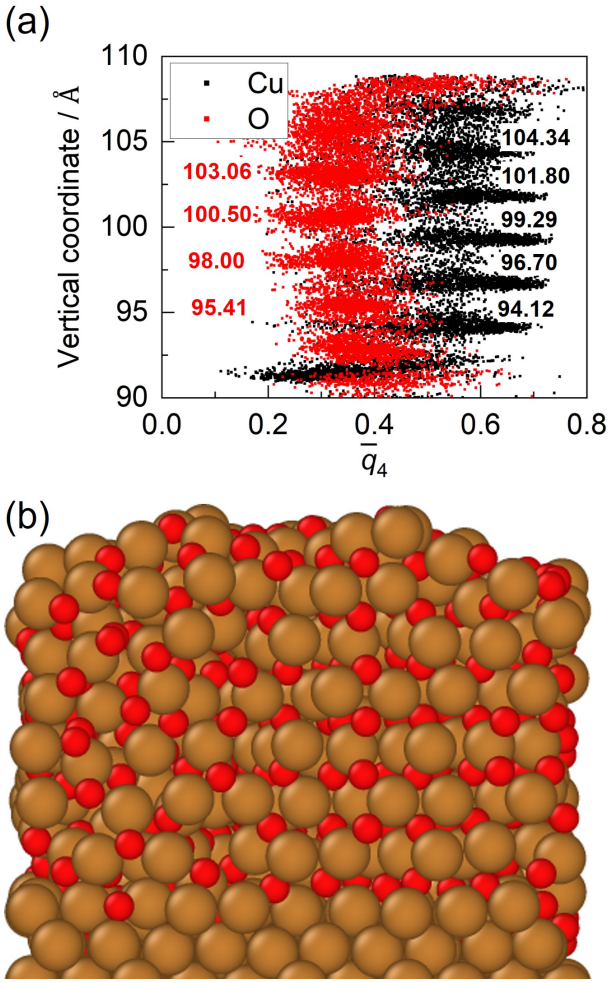


Figure 6.8: Bond orientational order parameter (BOO) on (10×10) CuO (100) slabs and the corresponding structure. At 1/4 of the surface area, the ion impact frequency were maintained at 10000 steps per ion. The neutral to ion ratio were kept at 10. This led to a 4-fold increase in effective ion and neutral flux. The surface temperature is still maintained successfully (*i.e.* surface has enough time to cool down before next ion impact. The analysis indicates clustering in the z-direction, in addition to in the BOO parameter values. Looking at the structure suggests the presence of crystalline CuO domains.

growth rate depends on the distribution of temperatures at different film thicknesses. To obtain this distribution, MD simulations used to obtain the depth of agitations in figure 6.4 were reused, since the deposition frequencies matched those in long-time MD (figure 6.2). The oxide layer is divided into 20 slices along the z-coordinate of the atoms, and the temperature distribution in each slice was calculated separately within each slice for each atom. Figure 6.9 provides an example. From this temperature distribution, the "per-particle rate" could be calculated once the apparent activation energy E_A and pre-exponential factor A are specified, as a function of film thickness. The "per-particle rates" are then averaged within each slice to obtain a thickness-rate relation. The rate is integrated over time to get the reproduced film thickness (equation (6.3)). Optimization is performed on the objective function in equation (6.4) over E_A and A to minimize the difference between the thickness of the reproduced film and the thickness of the MD simulated film, to obtain the optimal E_A and the prefactor A,

$$\frac{dL}{dt} = A \cdot \exp\left(-\frac{E_A}{k_B T}\right) \tag{6.1}$$

$$\bar{r}(L) = \frac{1}{N}\sum_{i}^{N} \frac{dL}{dt} \tag{6.2}$$

$$L^{\mathrm{rp}}(t_i) = L^{\mathrm{rp}}(t_{i-1}) + \bar{r}(L^{\mathrm{rp}}(t_{i-1}))\delta t \tag{6.3}$$

$$F = \sqrt{\sum_{t=t_0,\ldots,t_i,\ldots,t_f} (L^{\mathrm{rp}}(t_i) - L^{\mathrm{md}}(t_i))^2} \tag{6.4}$$

Temperature distributions in the film clearly depends on the kinetic energy of the impact ions, the atomic configurations of the slab, and the thickness. The subject slab is taken at the end of the long-time MD with the $T_s = 353\,\mathrm{K}$, $E_K = 20\,\mathrm{eV}$, n/i = 20, because this structure reached the greatest thickness. The structure was re-equilibrated at 353 K for a long time to desorb all the physisorbed O_2 molecules. In calculating the temperature response of the oxide slab to ion impact, a single ion was launched at the surface at 500 different random positions above the slab, and the trajectory is followed in NVE ensemble for 10000 steps, corresponding to roughly 4 ps. Two sets of simulations were performed with $E_k = 10\,\mathrm{eV}$ and $E_k = 20\,\mathrm{eV}$. The oxide region in the resulting trajectories (500 for one E_K) was divided into 20 layers in the z-direction, and the distributions of atomic kinetic energy

were calculated and converted to the distribution of "atomic temperatures" based on the equipartition principle.

The procedure is illustrated in figures 6.9 to 6.11. In figure 6.9, the temperature distribution for different depths down the oxide film is visualized in a scatter plot. Due to the enormous number of atoms, only points corresponding to atoms with temperatures above 500 K are shown. One can clearly observe that "hot" atoms are limited to a short range. In figure 6.10, the distribution at a fixed z-coordinate window is plotted as two histogram for $E_K = 10\,\text{eV}$ and $E_K = 20\,\text{eV}$, showing a large increase in the count of hot atoms as ion kinetic energy is increased. In figure 6.11, the averaged rates are plotted as a function of depths (z) for a set of activation energies. The scale is arbitrary as the pre-exponential factor can be used to scale the rates uniformly. The effect of temperature is clearly seen by the higher rates at higher z-coordinate (closer to the oxide/plasma interface). This rate-thickness relation is then integrated from a chosen starting point on the MD-derived thickness-time curve, following the shooting method for a initial value problem: at given thickness, the growth rate is found from the figure 6.11, and the thickness is incremented by $\frac{dz}{dt}\delta t$ where δt is a small integration time step.

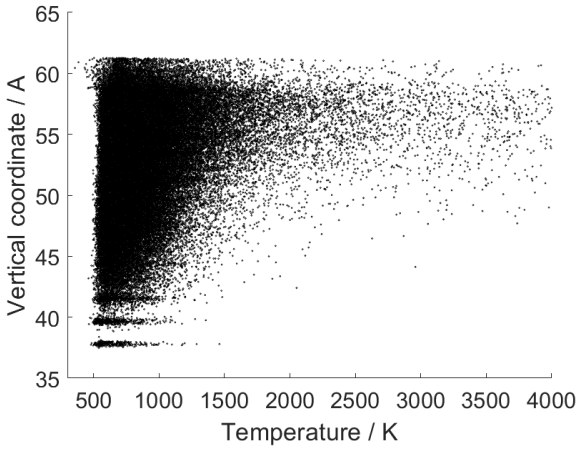


Figure 6.9: Temperature distribution with 10eV ion impacts

figure 6.12 shows the result of this optimization. Fitting E_A and A to the growth curve of $T_s = 353\,\text{K}$, $E_K = 10\,\text{eV}$, n/i = 10 gives an apparent activation energy of 1.0 eV. The fit starts at $t_0 > 0$ to skip the early growth regime dominated by the neutral-to-ion ratio, which

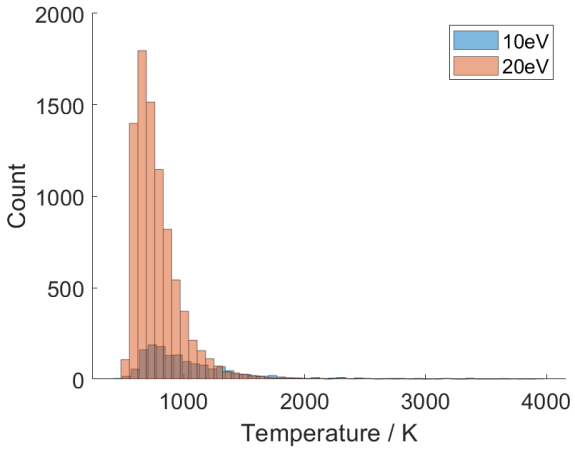


Figure 6.10: Temperature distribution at fixed vertical coordinate $z = 52$, near the middle of the oxide film

is not captured by the simple model. Long-time behavior is emphasized in equation (6.4) due to the larger number of points in this region, leading to the reproduced curve (figure 6.12, dashed black curve) lying close to the MD curve (solid black curve) at 10 eV for simulation times beyond 3 ns. The discrepancy at smaller oxide thicknesses cannot be eliminated by adjusting the starting point of the fit (t_0) or improving the fitting process. This could indicate more than one activated process with distinct kinetic parameters, which is reasonable given the different film characteristics at different depths, as observed with BOO parameters and the atomic concentrations. It is possible that the transport of oxygen species is ballistic in the porous, low-density region near the plasma-oxide interface and activated-diffusional near the high-density oxide-metal interface. Another explanation for the initial mismatch could be the unrealistic zeroth-order assumption in the form of the rate law. As shown next, the existence of a chemical driving force suggests that the oxidation process is not of the zeroth order even for thick oxides. At the long-time limit, the growth rates are well captured by the fitting. Note that the kinetic parameters fitted to the temperature distribution from 10 eV ion impacts can be used to directly predict the growth curves of 20 eV ion impacts(figure 6.12, red curves), suggesting that the rate-limiting process is independent of the kinetic energy of the ion.

The fitted activation barrier can be compared to a direct calculation of the diffusion

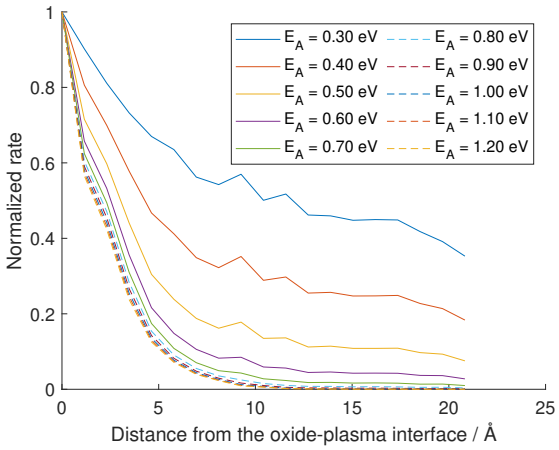


Figure 6.11: Normalized oxide growth rates as functions of distance downward from the oxide-plasma interface. Rates at each apparent activation energy is normalized by their maximum rates.

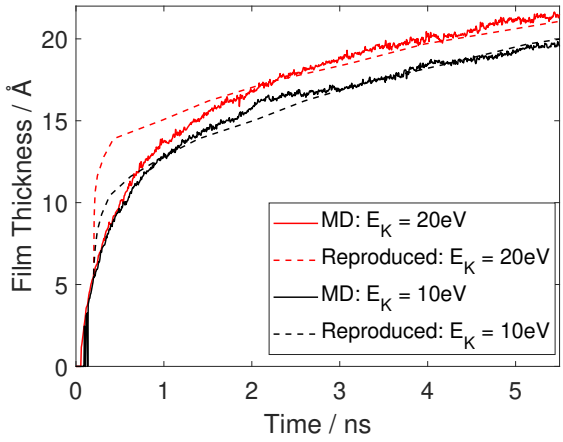


Figure 6.12: Comparison of the fitted thickness curve to the MD result. The apparent activation energy E_A and pre-exponential factor A is fitted to the thickness-time relation of the $T_s = 353\,\mathrm{K}, E_k = 10\,\mathrm{eV}, n/i = 10$ trajectory from the temperature distribution obtained using $E_k = 10\,\mathrm{eV}$ ions. The resulting optimal values of $E_A = 1.01\,\mathrm{eV}$ and $A = 5.63 \times 10^7\,\mathrm{nm/ns}$ are used to reproduce the $T_s = 353\,\mathrm{K}, E_k = 10\,\mathrm{eV}, n/i = 10$ and $T_s = 353\,\mathrm{K}, E_k = 20\,\mathrm{eV}, n/i = 10$, using temperature distribution obtained using $E_k = 10\,\mathrm{eV}$ and $E_k = 20\,\mathrm{eV}$, respectively.

barriers in well-defined lattices. The activation energy of oxygen diffusion in crystalline copper has been experimentally reported to be in the range of 0.6 eV to 0.7 eV,[11,100] and computationally calculated to be around 0.5 eV.[140] The fitted result indicates that the plasma oxidation process is quite different from the diffusion of oxygen atom impurities in copper lattice. On the other hand, the barrier to oxygen vacancy diffusion in CuO is calculated to be 1.11 eV,[169] suggesting that O vacancy diffusion could be a mediating mechanism for oxide growth, although a closer atomic level analysis of the trajectories is needed to fully establish this. To put the fitted barrier in context, we have calculated the barrier of oxygen interstitial diffusion in copper and oxygen vacancy diffusion in monoclinic CuO with our potential. The results are shown in Figure 6.13. Nudged elastic band (NEB) method with 9 images were used in these calculations.

The barrier to oxygen interstitial diffusion in Cu lies near 0.5 eV, similar to results calculated using PW91 functional.[140] These calculations are performed with bulk Cu lattices obtained directly from materials project (mp-30). The conventional cell is repeated 4 times in each direction, and the initial and final positions are neighboring octahedral and tetrahedral interstitial sites, respectively.

Similarly, for the CuO oxygen vacancy diffusion barrier, the lattice is obtained from materials project (mp-704645).The paths selected follows the notation used in Li et. al. [169]. The barrier to oxygen vacancy diffusion along the three paths are calculated to be 1.0 eV, 1.0 eV, and 1.8 eV, respectively for paths 1, 2, and 3, all of which lies very close to the literature DFT values calculated using the PBE functional (1.11 eV, 1.11 eV, and 1.89 eV, respectively).

The internal structure of the modeled oxide film can be considered to have two parts. The atomic concentrations of the O and Cu atoms are shown in figure 6.14 as a function of thickness. It suggests that the film has a porous region within $\sim$ 1 nm below the oxide-plasma surface, indicated by the lower overall density, and a dense region beyond. In the dense region, a gradient of oxygen and copper concentration can be seen. The observed gradient is not a result of an uneven oxidation front: The onset of the gradient, at $\sim$ 8 Å above the oxidation front (as defined in appendix C.1.1), does not have local environments

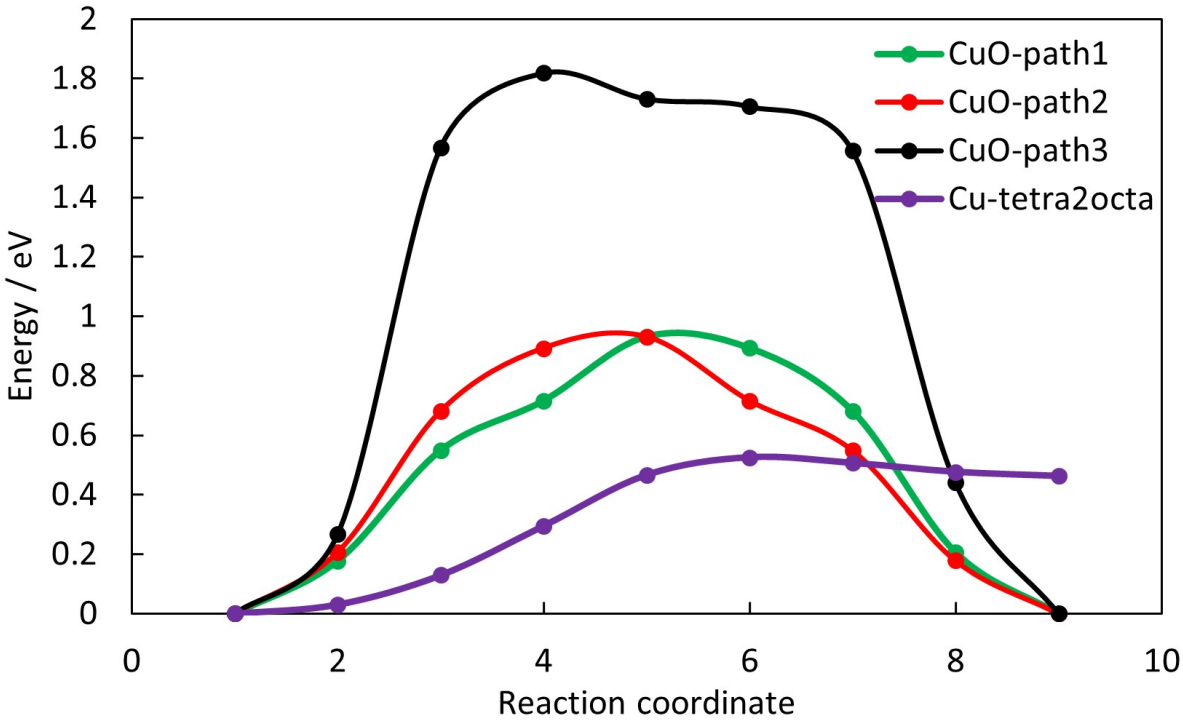


Figure 6.13: Barrier obtained using nudged-elastic band method with the trained NNP. The lines labelled with CuO refer to oxygen vacancy diffusion in CuO, and the paths refer to Li et. al. [169]. The line labelled with Cu refer to oxygen interstitial diffusion in Cu.

of metallic copper, as indicated by the bond orientation order parameters. The existence of such a concentration gradient (and thus the chemical potential gradient) indicates a chemical driving force behind the growth phenomena observed collectively in figure 6.3. It also explains the trend observed when the neutrals are removed from the simulation (figure 6.4c), where the growth seems to follow the case with both ions and neutrals for $\sim 2\,\text{ns}$, but eventually becomes slower. Without the frequent addition of neutral oxygen species to the plasma-oxide interface, the infrequent ion addition alone cannot sustain the same concentration gradient for long. Once the driving force of the chemical potential is gone, the growth slows. The time lag between stopping the neutrals and significant slowdown of the growth is related to the characteristic mass diffusion time scale in the oxide film, which is certainly much greater than the thermal diffusion time scale. Recalling the temperature gradient discussed earlier, the plasma oxidation shows an ion-neutral synergy similar to that observed in reactive ion

etch (RIE). Namely, two gradients are required to sustain fast growth: the chemical potential gradient provided by the neutrals, and energy gradient provided by the ions. This explains the effect of stopping ions or neutrals in figure 6.4c.

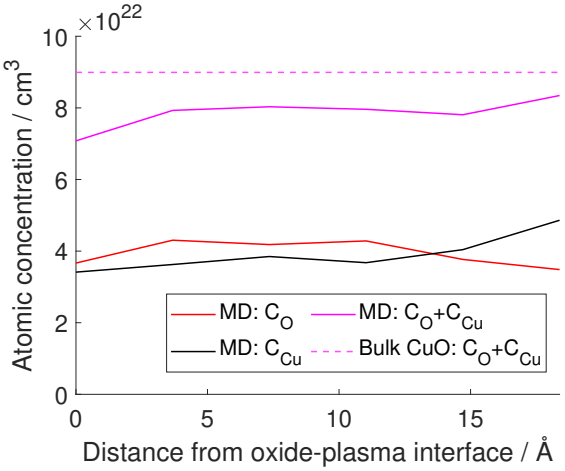


Figure 6.14: Atomic concentration in the oxide film. The film is taken from the $T_s = 353\,\mathrm{K}, E_k = 20\,\mathrm{eV}, n/i = 10$ trajectory at time $t = 5.54\,\mathrm{ns}$. The result here indicate a porous region where the total atomic density is lower, a dense region where the Cu and O concentration is nearly constant, a region where a concentration gradient of O is observed.

Some clarification is due in relation to a recent contribution by Kunze et. al.,[168] where *in situ* observation of plasma oxidation of single-crystal copper metal at room temperature was reported. In their experiments on Cu (100), only Cu_2O was observed initially, and only prolonged exposure leads to a "sandwich" structure of CuO, Cu_2O and Cu. This is consistent with *ex situ* XPS results under conditions targeted by our model showing the existence of CuO.[173] The "sandwich" structure is not observed in the current MD simulation, likely due to the much lower thickness that we were able to reach in the MD simulation: while a concentration gradient is observed in our MD structures, a Cu_2O-like environment is not observed via the BOO parameters. The molecular dynamics technique also does not provide electronic structure information, preventing a direct comparison with the auger electron spectroscopy (AES) and near-edge X-ray absorption fine structure (NEXAFS) characterizations. Fundamentally, our MD simulation only probes the very initial, transient stage of the oxidation process, on the order of nanoseconds to microseconds. To be directly comparable,

the notorious time scale gap in MD needs to be bridged.

Important differences exist between the work of Kunze *et. al.* and the target experimental process in this work. First, the pressures used in Kunze*et. al.* are at least two orders of magnitude lower than the experimental process targeted by our model. For thermal oxidation, higher pressure and longer exposure times favor the formation of CuO over Cu_2O, and lead to a similar "sandwich" structure at intermediate temperatures.[135] An analogous phenomenon seems to be present when comparing the LEEM (low energy electron microscopy) result with the NEXAFS result in Kunze *et. al.*, conducted at different pressures. It is possible that increasing the pressure by 2 more orders of magnitude, as in our target process, further diminishs the fraction of Cu_2O. Second, the differences in substrate bias (400V DC versus unbiased) and the plasma source (electron cyclotron resonance, ECR *vs.* inductive coupled plasma, ICP) could lead to different plasma densities and ion energy distributions. These considerations highlight the complexity of plasma reactions and the need for further work.

A strategy is proposed to achieve a self-limited oxidation layer in plasma processing. Because oxidation beyond the first few layers, as demonstrated here, is mostly a diffusion-limited process, to contain a diffusion process to the surface, the substrate temperature itself should be lowered to the point where diffusion practically stops during the processing time window. The thickness is then tuned by varying the plasma power, which affects the kinetic energy of the ions and, in turn, the penetration thickness of the thermal spike. The thickness of the oxide created would be self-limiting in the sense that it depends only on the kinetic energy but not on the exposure time. This can lead to a feasible ALE process if the temperature can be cycled between the steps, since the subsequent removal step requires in general higher temperatures to facilitate etching product desorption. Increasing the plasma power is expected to give films with higher crystallinity since the impact frequency is higher. While this is beneficial for self-limiting the oxidation, since a more crystalline film is shown to slow oxidation down in this study, crystalline films may be more difficult to etch for the subsequent removal step.[174]

To ensure the results is not particular to the Cu (100) slab, similar long-time MD simula-

tions were performed on the Cu (111) surface. Figure C.3 shows the evolution of the chosen rectangular unit cell on the Cu(111) surface, and figure C.2 shows the thickness and oxygen content as functions of simulation time. Since the noise is considerably higher due to the smaller number of atoms per layer, the analysis of thermal spike penetration, local concentration gradient, *etc.* is not performed. The surface undergoes similar stages of oxidation. The thickness and oxygen incorporation curves show similar kinetic energy controlled behavior, as on the (100) surface(see figure C.2). Since the kinetic energy of the ions is much higher than the energies of Cu-Cu and CuO bonds, both surfaces should respond similarly. Since (111) and (100) are the low-energy surfaces exposed in equilibrium conditions, qualitatively the conclusions should hold on practical polycrystalline Cu surfaces. Quantitatively, the close-packed [111] direction is expected to slow down diffusion.

6.4 Conclusions

The results demonstrate that the oxidation under reported experimental conditions is not self-limiting, in agreement with the experimental observations. The oxide film is shown to be porous and amorphous CuO on top of a potentially crystalline CuO. Whether a crystalline CuO phase is formed depends sensitively on the ion flux and neutral flux used in the simulation. No Cu_2O can be identified from the MD simulations under our chosen conditions. The oxidation trajectories are explained in terms of an interplay between thermal fluctuation at nominal process temperature and transient, intense thermal agitations resulting from ion impacts. Namely, higher energy ions deliver thermal energy deeper into the oxide film, accelerating diffusion/interface reaction in the substrate. Ion impact is found to create a non-equilibrium "temperature gradient" inside the oxide film, which is subsequently fitted to the oxide growth rate to obtain an apparent activation energy of 1.0 eV, $\sim 30\%$ higher than that of oxygen diffusion in Cu. On the other hand, an abundance of the reactive neutral species is also important because they provide a chemical potential gradient through the film, driving the oxidation front forward. The immediate consequences for creating a self-limited oxidation process is that low-temperature substrate should be used whenever possible to

limit the uncontrolled vertical oxidation. The thickness and potential film morphology at the atomic scale could then be controlled via plasma power.

CHAPTER 7

Conclusion

In this dissertation, we have demonstrated the importance of atomistic understanding in designing new plasma-thermal atomic layer etching (ALE) processes on metals and improving existing processes. The combination of experimental and theoretical insights has led to the discovery of a nitrogen plasma-based ALE process on metals, which exhibits similar selectivity and directionality as the oxygen plasma-based process, as predicted theoretically. Moreover, efforts are underway to validate the theoretical prediction of self-limited copper oxidation at low temperatures.

Equally important, the theoretical tools developed in this dissertation can pave the way for further discoveries. By utilizing the thermodynamic model, we have already discovered (and experimentally confirmed) that a nitridation-based plasma yields similar etching characteristics. The application of the free energy model, along with techniques for systematic and grand-canonical exploration of potential energy landscapes, has seen significant improvements in recent years. Combining these approaches could lead to rapid process screening with impacts extending beyond metal ALE, as the model is not restricted to metallic systems.

The copper-oxygen interatomic potential is a general-purpose machine learning potential capable of describing high-energy plasma-surface interactions. Its first-principles level description of interatomic interactions allowed us to draw confident conclusions about the growth dynamics, including the effects of substrate temperature, ion kinetic energy, and more. The development of this potential has shed light on fundamental difficulties associated with machine learning potential development. Presently, work is underway to develop a similar potential for the nitrogen-plasma based process, and we are leveraging the experience gained from the oxidation potential development to automate the construction of binary

interaction potentials as much as possible.

The simulation protocol for plasma oxidation of copper has undergone significant testing, and we have obtained best practices regarding simulation parameters and computational setup to efficiently perform such large-scale simulations. We were able to demonstrate that energetic ions play a substantial role in promoting oxide growth within their range of influence, with the radius of influence strongly dependent on the kinetic energy. Furthermore, the effect of substrate temperature has been shown to act as an isotropic heat source, driving the oxygen profile towards increased isotropy. As the directionality of plasma-thermal ALE solely derives from the plasma step, it also affects the directionality of the entire process.

Having an atomically-resolved oxidation trajectory opens up a fresh domain of characterization. For the first time, we can understand the film morphology, chemical composition, local environment, etc., at the atomistic level. This thesis has included a few such characterizations. Additionally, beyond simply calculating the growth rate, we have shown that the local atomic environment of the amorphous oxide film derived from the plasma closely resembles that of bulk CuO. Moreover, we demonstrated that repeated "micro-annealing" steps lead to the precipitation of crystalline domains from the amorphous CuO domains. Such observations have significant implications for modeling etching reactions, and more broadly, they demonstrate the power of machine-learning potentials to capture potential energy surfaces with unprecedented accuracy.

Appendix A

Supplementary information for the thermodynamic model

A.1 Estimation of reactant and product pressures

The reactant and product pressures are estimated roughly based on the fluxes used in the experiments. Assuming complete vaporization and perfect mixing, the etchant pressure is calculated by:

$$P_{\text{etchant}} = \frac{f_{\text{etchant}}}{f_{\text{etchant}} + f_{\text{carrier}}} P_{\text{total}} \tag{A.1}$$

Where f's are the molar flow rates of etchant and the carrier gas. It is reported in Ref.[159] that the flow rate of N_2 carrier gas is 100 sccm, the flow rate of formic acid into the vaporizer is 0.1 g/min, and $P_{\text{total}} = 350\,\text{Torr}$. Thus $f_{\text{etchant}} = 0.0021\,\text{mol/min}$, and $f_{\text{carrier}} = 0.0035\,\text{mol/min}$. This gives $P_{\text{etchant}} = 18\,025\,\text{Pa}$. The estimation of the product pressure is difficult and we simply put it as 10% of the etchant pressure. The proposed processes using formamidine as the etchant are simulated at the same conditions. We expect that the pressure of the product does not significantly affect thermodynamics since at such low pressures, the deviation from ideal gas can be neglected. We verified that alternative estimates of product pressure to be 5% and 20% of the etchant pressure only changes the removal step energies by around 0.02 eV, and does not change our conclusions.

A.2 Validation of Computational Methods

The relaxed crystal structure and their atomization energies are compared to experimental values in table A.1. The free energy is used throughout our study. The enthalpies are provided since the experimental free energy of formation of Cu_3N is not available. The results indicate that except for NiO, the calculated energies agree very well with experimental energies. The calculated values for the oxides are shown but not used since they are known pathological cases for the PBE exchange-correlation functional. The functional is retained nevertheless due to the need to calculate and compare energies on diverse systems, that include surfaces, bulk structures, organic molecules, and organometallic complexes.

Table A.1: Optimized bulk lattice and formation energy. All experimental values are obtained from *CRC Handbook of Chemistry and Physics*[115] unless otherwise stated. The two numbers in Ni_3N lattice correspond to a and c, respectively. The calculated values are obtained via self-consistent minimization in VASP with the PBE exchange-correlation functional. The values used are marked with boldface.

	Lattice [Å]		Cohesive enthalpy [eV]		Cohesive free energy [eV]	
System	Exp.	Calc.	Exp.	Calc.	Exp.	Calc.
Ni	3.52	3.52	4.45	4.60	3.98	**4.19**
Cu	3.62	3.64	3.50	3.44	3.09	**3.05**
NiO	4.18	4.20	9.52[1]	8.82	**8.58**	7.78
Cu_2O	4.27	4.31	11.32	11.46	**10.09**	10.03
Ni_3N	4.64/4.31[61]	4.62/4.31	N/A	18.75	N/A	**17.14**
Cu_3N	3.81	3.83	14.15[25]	14.25	N/A	**12.71**

Van der Waals interaction could have a potentially important influence on our removal step energy results. The contribution of van der Waals interaction to the removal step energy is not quantified directly, since the PBE+dDsC[92,93] formalism does not accurately reproduce the bulk formation energies, which is a significant term in equation (3.5). Alternative methods, including the vdw-DF family of functionals, suffer similar problem..[121]

Instead, we chose to quantify the contribution by comparing the adsorption energies with and without this contribution. Figure A.1 shows the adsorption energies (equation (A.2)) calculated with and without the dDsC contributions on the Cu (100) surface and Ni (111) surface, with N and O adsorbed. With the dDsC correction, the adsorption becomes more favorable by ~ 0.1 eV on both Cu and Ni systems. Notably, the magnitude of the correction does not change much with increasing coverage. This uniformity over the entire range of coverage investigated suggest that the deviation introduced by neglecting van der Waals interaction is approximately 0.05 eV on the Cu systems and 0.1 eV on the Ni systems. Errors of this magnitude influences the quantitative thermochemical predictions, but they do not invalidate our results of whether given etching chemistry leads to favorable reaction.

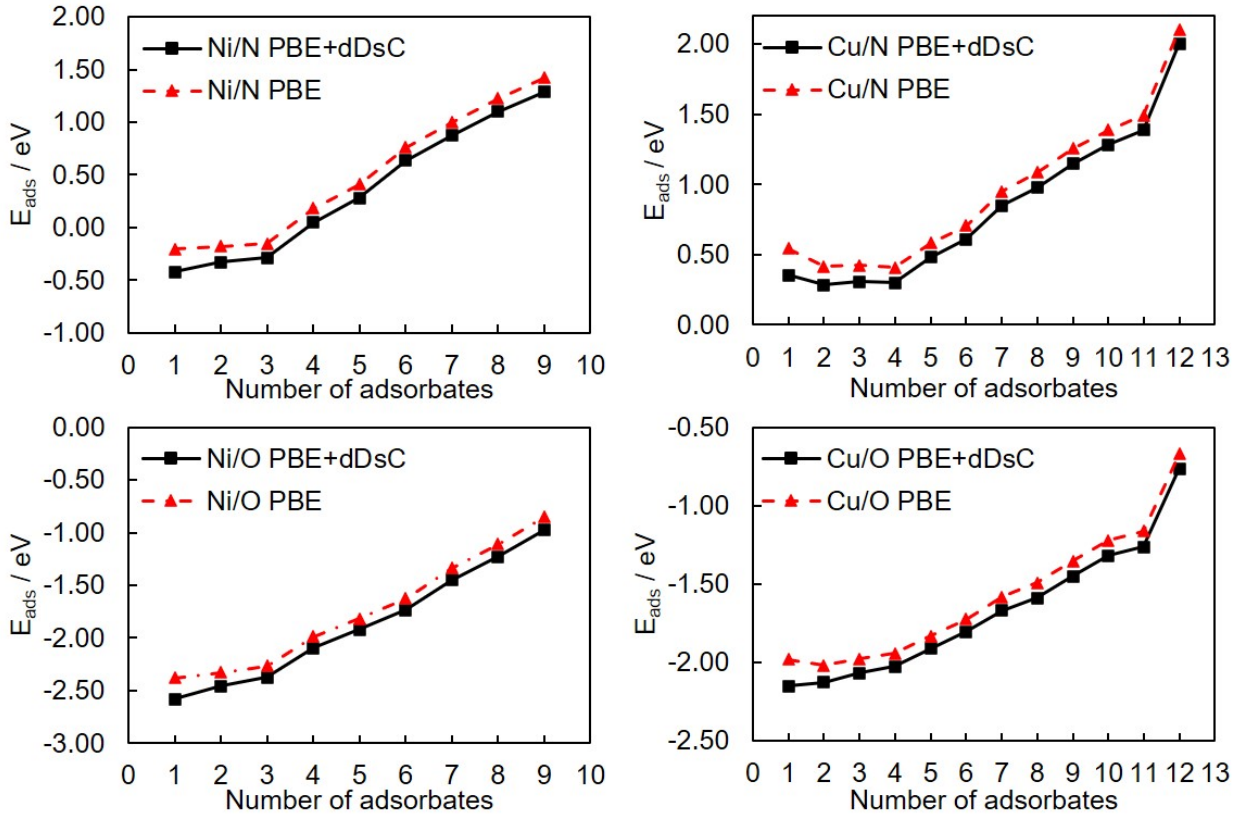


Figure A.1: Adsorption energies with and without dDsC correction

A.3 Adsorption configuration search setup

figure A.2 shows the supercell used in the configuration search and the "symmetry unique" sites. Except for the sites that merged into the same post-relaxation structure, the single-adsorbate adsorption energies of the four combinations of metals (Ni and Cu) and modifiers (N and O) are shown in table A.2, calculated from equation (A.2) (written for Ni/O system).

$$\Delta G_{\mathrm{ads}} = (G_{\mathrm{Ni_s|O_y}} - G_{\mathrm{Ni_s|}} - y\mu_{\mathrm{O}}) \approx (E_{\mathrm{Ni_s|O_y}} - E_{\mathrm{Ni_s|}} - y\mu_{\mathrm{O}})/y \tag{A.2}$$

Table A.2: Adsorption sites, labels, and adsorption energies (in eV) referenced to O_2 and N_2 at 80°C, 350 Torr. A blank cell means the site is not stable in relaxation and coalesced to other sites.

			ΔE_{ads}[eV]			
Termination	Site name	Site label	Ni/N	Ni/O	Cu/N	Cu/O
(100)	t	A1	2.43	−0.43	3.58	0.20
	b	A2	0.78	−1.63	2.15	−1.01
	4h	A3	−0.80	−2.34	0.56	−1.85
(110)	t-tr	B5	0.24	−1.53	1.46	−1.09
	t-pk	B4	2.42	−0.58	3.96	0.43
	sb	B2	0.81	−1.76	2.37	−0.87
	lb	B3	−0.05	−1.54	1.52	−1.14
	3h	B1	0.29	−1.80		−1.21
(111)	t	C3	2.58	−0.22	3.98	0.70
	3h-hcp	C2	0.11	−1.92	1.93	−1.00
	3h-fcc	C1	0.08	−2.04	1.85	−1.09
(210)	t-tr	D6			2.23	−0.81
	t-pk	D1	2.41	−0.70		
	b-tr	D2	0.31	−1.59	1.88	−0.81
	b-pk	D7	−0.48	−2.26	1.14	−1.55
	4h	D3	−0.06	−1.66	1.42	−1.20
	3h-1	D4	0.34	−1.67	1.42	−1.12
	3h-2	D5	0.34	−1.67		

…continued

Termination	Site name	ΔE_{ads} [eV]				
		Site label	Ni/N	Ni/O	Cu/N	Cu/O
(211)	t-st	E9	2.70			
	t-pk	E1	2.46	−0.59		
	b-pk	E2		−1.86		
	b-cs	E3	1.16	−1.22	2.35	−0.60
	4h	E4	−0.24	−1.82	0.98	−1.39
	3h-hcp-st-l	E6	0.28	−1.68	1.85	−1.07
	3h-hcp-st-h	E8	−0.04	−2.14	1.48	−1.45
	3h-fcc-st-l	E5	0.29	−1.62		−1.14
	3h-fcc-st-h	E7	0.24	−1.92	1.58	−1.30
(221)	t-tr	F7	0.16	−0.58	1.29	
	t-pk	F1	2.45		3.80	
	lb	F6	−0.09	−1.56	1.40	−1.07
	b-pk	F2		−1.82		
	3h-hcp-st-l	F5	0.20	−1.83	1.74	−1.16
	3h-hcp-st-h	F3	0.13	−1.84	1.67	−1.17
	3h-fcc-st-m	F4	0.14	−1.91	1.61	−1.34
	3h-fcc-st-l	F10		−1.58		−0.98
	3h-fcc-st-h	F8	0.00	−2.18	1.39	−1.60
	3h-hcp-cs	F9	0.26	−1.85		−1.38
(311)	b-tr	G2	1.17			−0.59
	b-pk	G1		−1.88	2.23	
	4h	G3	−0.20	−1.80	1.12	−1.36
	3h-hcp	G4	0.08	−1.99	1.61	−1.33
	3h-fcc	G5	0.41	−1.61		−1.07

For validation against experiments, the adsorption energies on selected experimentally observed low-coverage surface adsorption structures are also calculated and shown in table A.3. Experimental structures that do not allow a direct comparison(i.e., those with extra or missing metal atoms) to the ones found through the configuration search are marked with an asterisk.

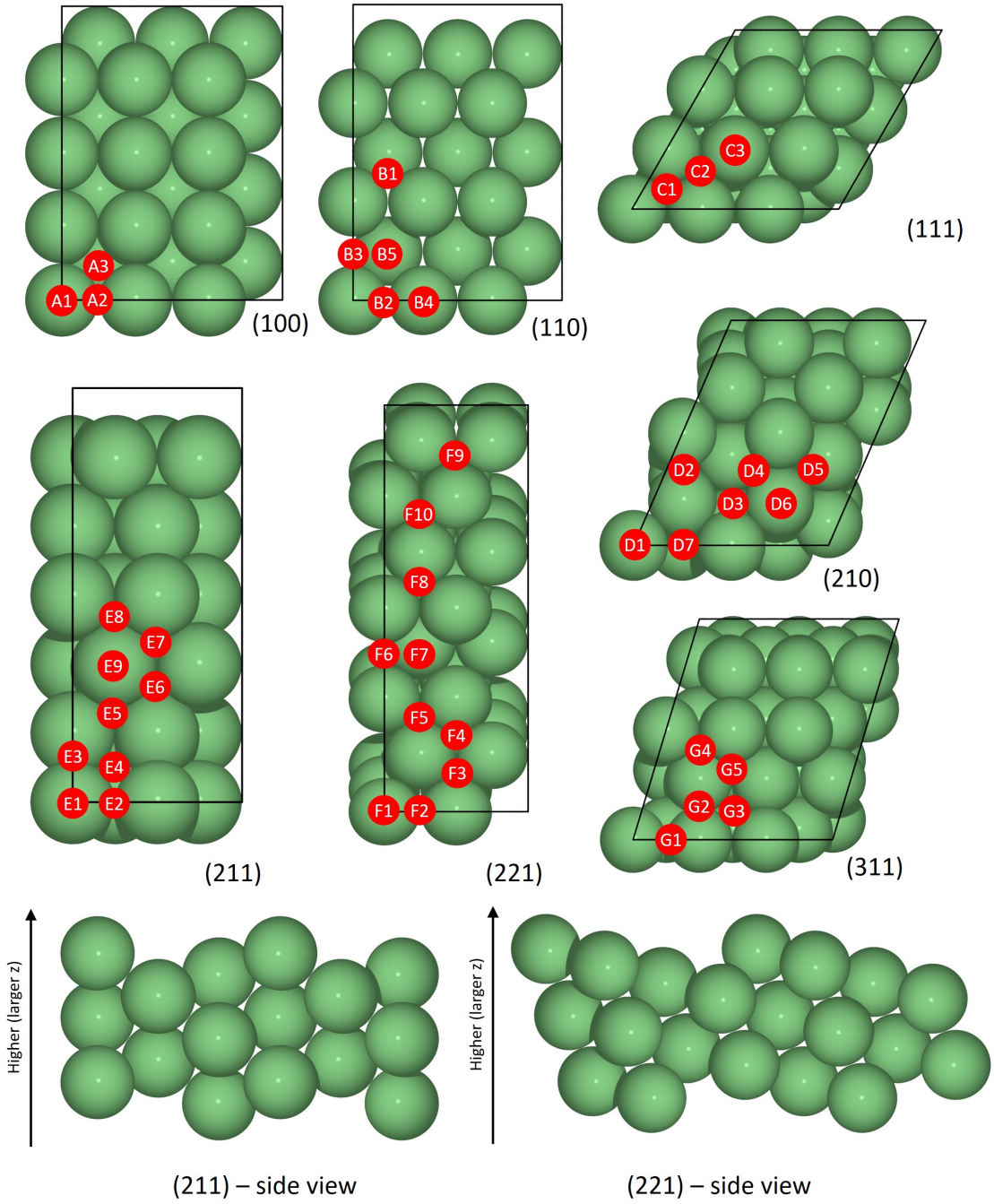


Figure A.2: Adsorption sites on (100), (111), (110), (210), (211), (221), and (311) terminations of Ni and Cu. The supercells correspond to the actual slab model used in the configuration search. The (211) and (221) are shown in side view in addition to top view. Site labels are explained in table A.2

Table A.3: Experimentally observed surface reconstructions and adsorption superstructures for the Ni/O and Cu/O systems. The adsorption energy values marked with * indicate that these structures do not have a well-defined reference 'bare' surface, as metal atoms are added/removed to these surfaces. The 'bare' surface term in equation (A.2) is approximately calculated from that of pristine surface, adjusted by the number of atoms. The Ni(110) missing row reconstruction and the Cu(110) pairing-row reconstruction had their bare surfaces recalculated, hence are not marked with *.

Substrate	Termination	θ_O [Å^{-2}]	Surface cell	Ref.	ΔG_{ads}[eV]
Ni	(100)	0.040	$p(2\times2)$	[10]	−1.92
Ni	(100)	0.081	$c(2\times2)$	[10]	−2.27
Ni	(110)	0.057	(2×1) missing-row	[53]	−2.48
Ni	(110)	0.076	(3×1) missing-row	[53]	−2.13
Ni	(111)	0.047	$p(2\times2)$	[9]	−1.52
Ni	(111)	0.062	$(\sqrt{3}\times\sqrt{3})R30°$	[9]	−1.60
Cu	(100)	0.076	$c(2\times2)$	[83]	−1.38
Cu	(100)	0.076	$(2\sqrt{2}\times\sqrt{2})R45°$	[83]	−1.38
Cu	(110)	0.054	(2×1) pairing-row	[83]	−1.80
Cu	(110)	0.071	$c(6\times2)$	[83]	−3.35 *
Cu	(111)	0.087	$p4$	[66]	−2.66 *
Cu	(111)	0.098	$p4 + O_F$	[66]	−0.77 *
Cu	(111)	0.076	$p4 + OCu_3$	[66]	−1.02 *

The supercells used in this study (see figure A.2) are selected and not thoroughly searched. Hence, it is possible that our configuration search cannot find more stable configurations that require a different repeating pattern. To quantify the magnitude of this error, figure A.3 compares the adsorption energies of structures from the configuration on different supercells. On (100), while the experimental adsorption structures of $p(2\times2)$, $c(2\times2)$ and $\sqrt{2}\times\sqrt{2}R45°$ is not representable by the supercell used in the main results, the deviation errors introduced is on the order of 0.1 eV. On (111) and (211), the alternative supercells, through the same configuration search algorithm, finds less stable structures than the supercell used. The results here suggest that the structures used in this study are close to the true minima at the coverages investigated. An error of 0.1 eV would not refute our conclusions.

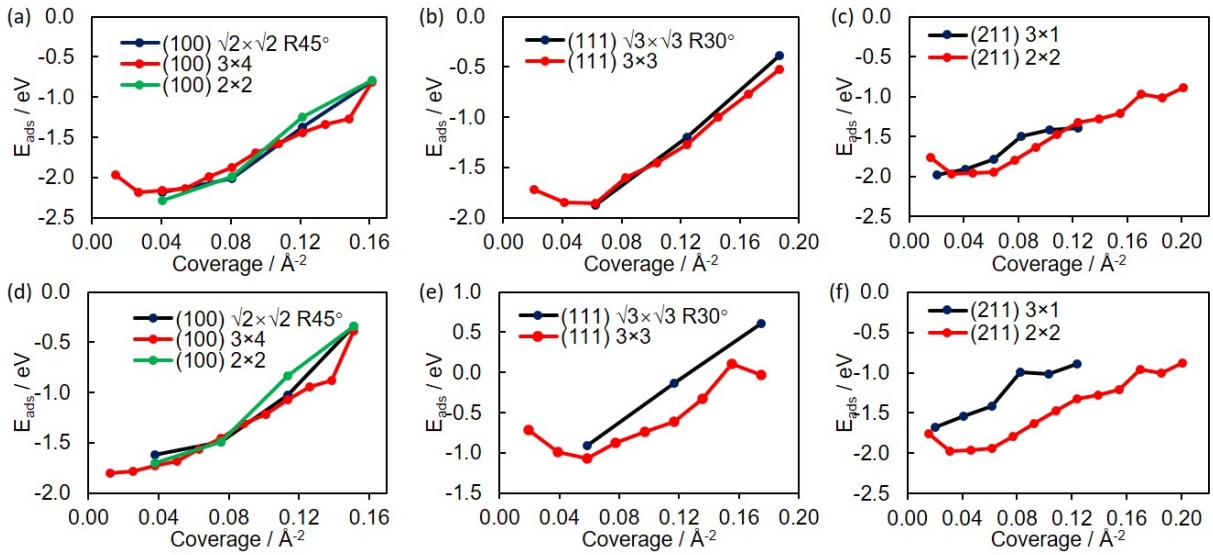


Figure A.3: Adsorption energies of structures found by configuration search on the Ni (100)(a), Ni(111)(b), Ni(211)(c), Cu(100)(d), Cu(111)(e), and Cu(211)(f) surfaces. The results from alternative supercells are marked with green and black lines, whereas the original supercells are marked with red lines. The adsorption energies are calculated with equation (A.2).

A.4 Averaging terminations using wulff construction

The chemical potential used to obtain the surface formation energies are shown in figure A.4. Its definition is given in equation (3.9). The surface energies resulting from the coverage-dependent chemical potential is shown in figure A.5. Its definition is given in equation (3.10). The fraction of surfaces given by the Wulff construction is shown in figure A.6. The use of these quantifies are explained in section 3.2.

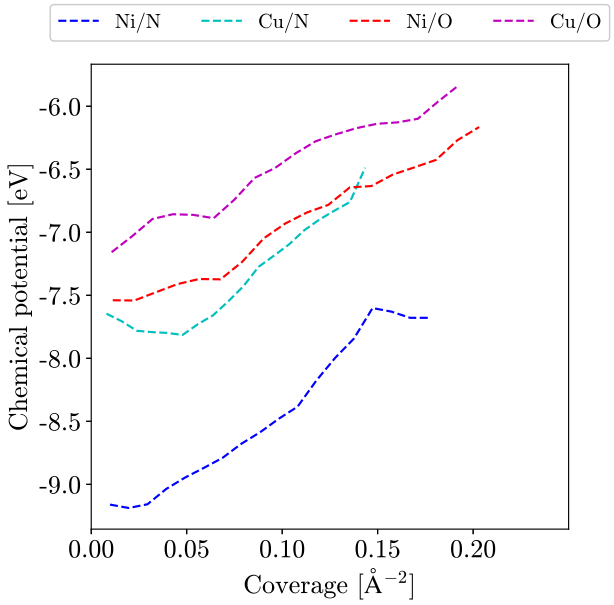


Figure A.4: Relation of the chemical potential to the surface coverage(area density of modifier atoms), obtained from equation (3.9).

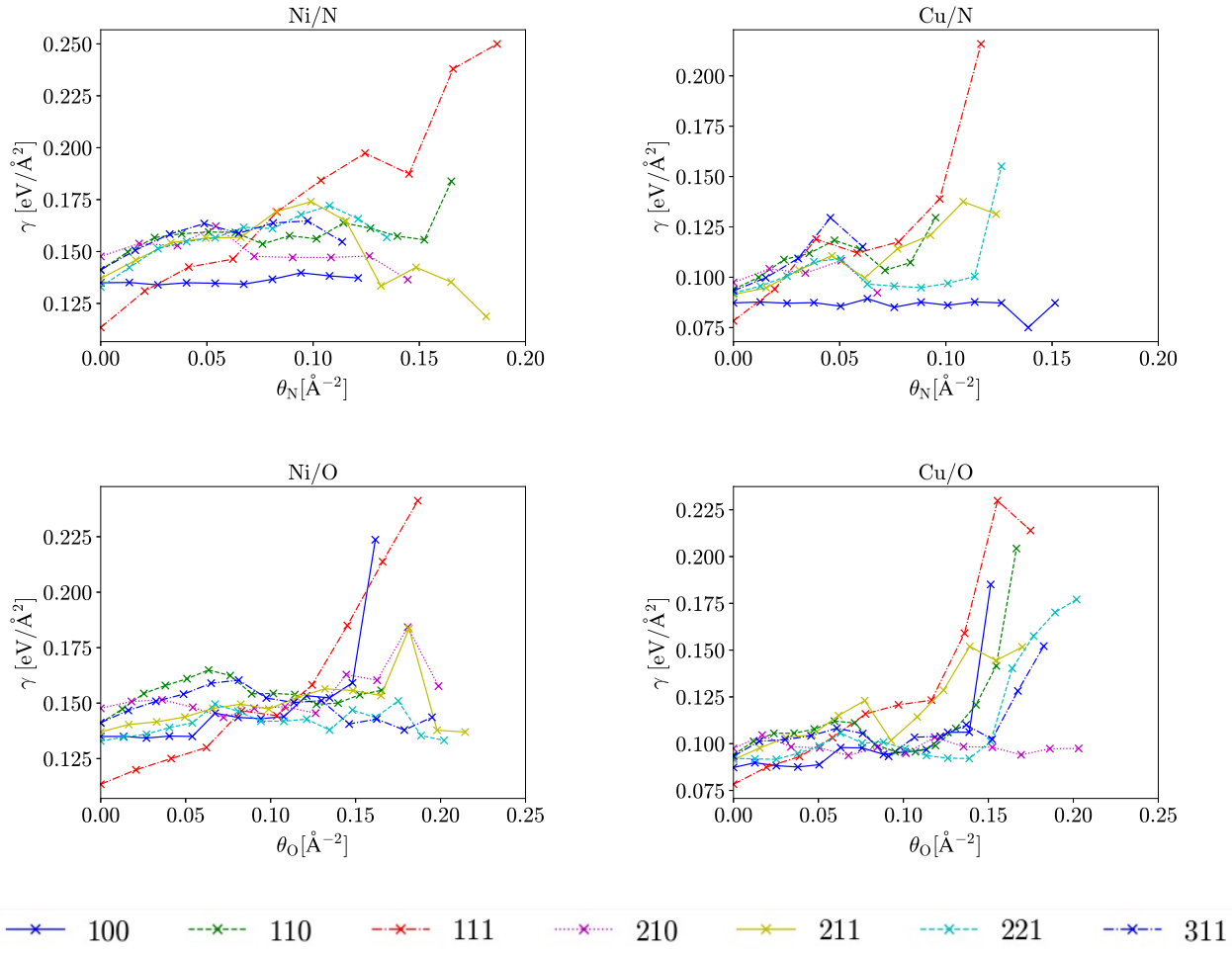


Figure A.5: Surface energies of terminations on the Ni/N, Cu/N, Ni/O and Cu/O systems as functions of coverage.

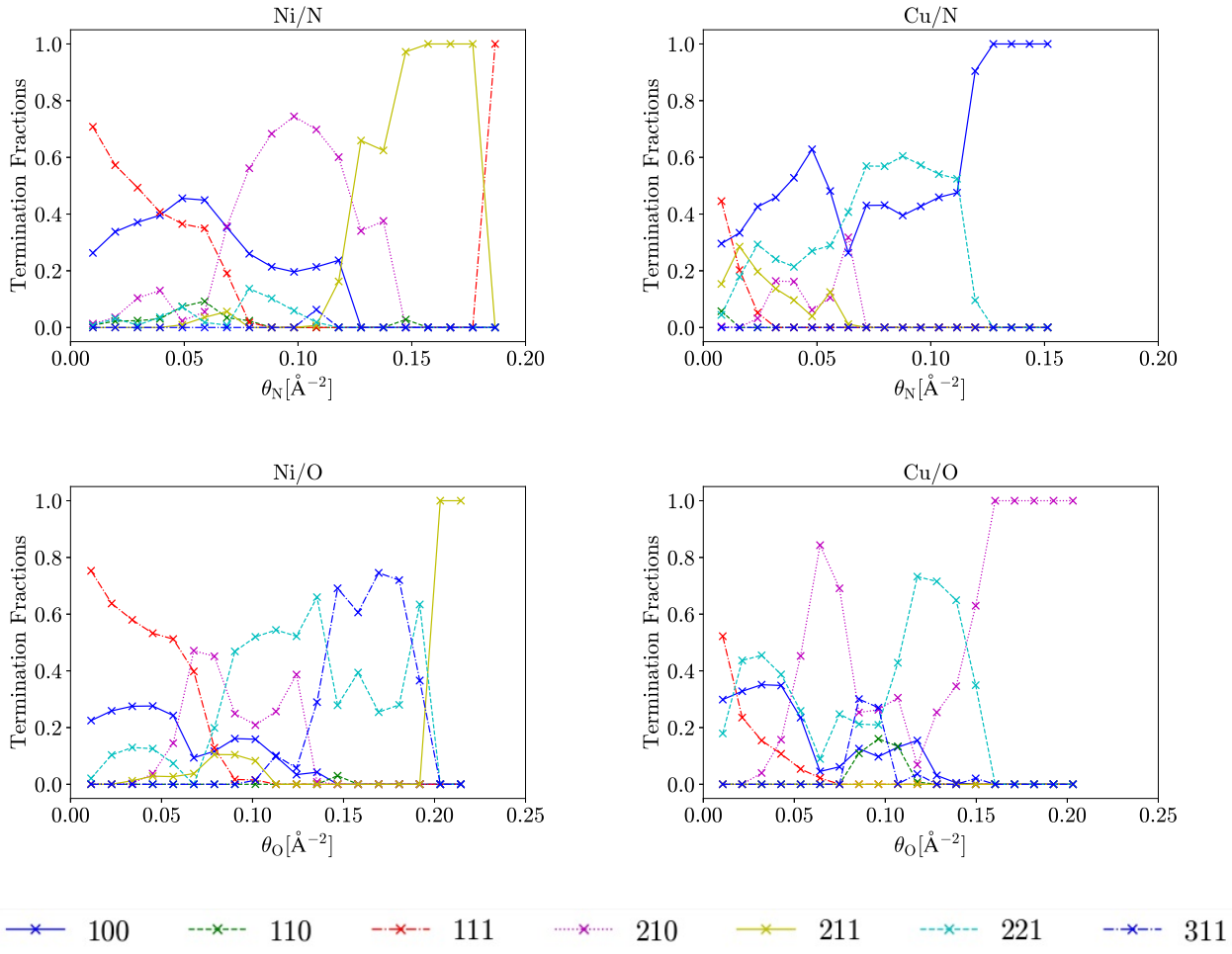


Figure A.6: Fractions of terminations on the Ni/N, Cu/N, Ni/O and Cu/O systems as functions of coverage, determined by Wulff construction.

A.5 Trajectories resulting from the configuration search

The relaxed structures at each step in the configuration search with the “greedy algorithm“ are shown in figure A.7 - figure A.10. The atomic coordinate files are provided as well. In the images, the green and brown substrate atoms represent Ni and Cu, respectively. The blue and red modifier atoms represent N and O atoms, respectively.

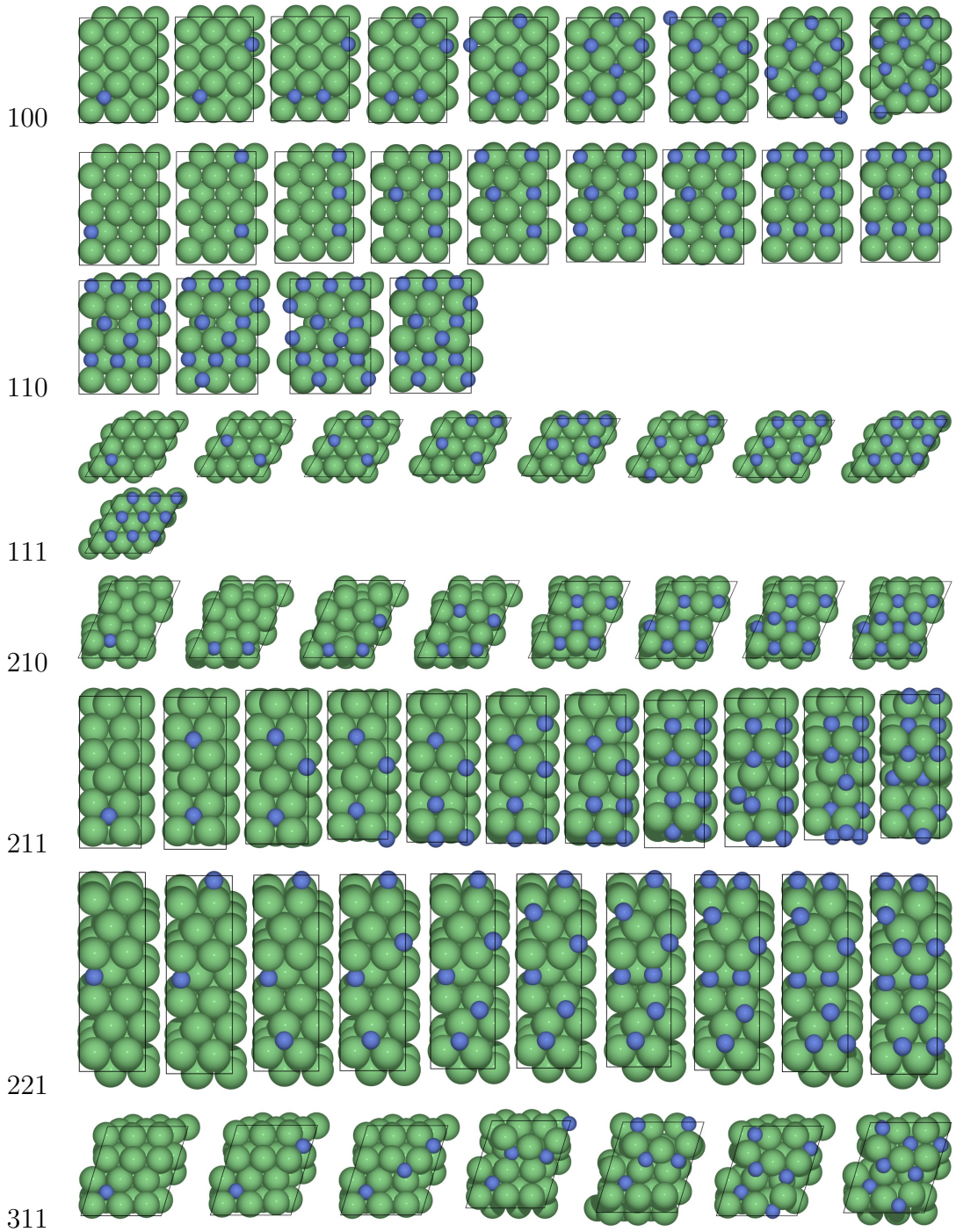


Figure A.7: Configuration search iterations for Ni/N system

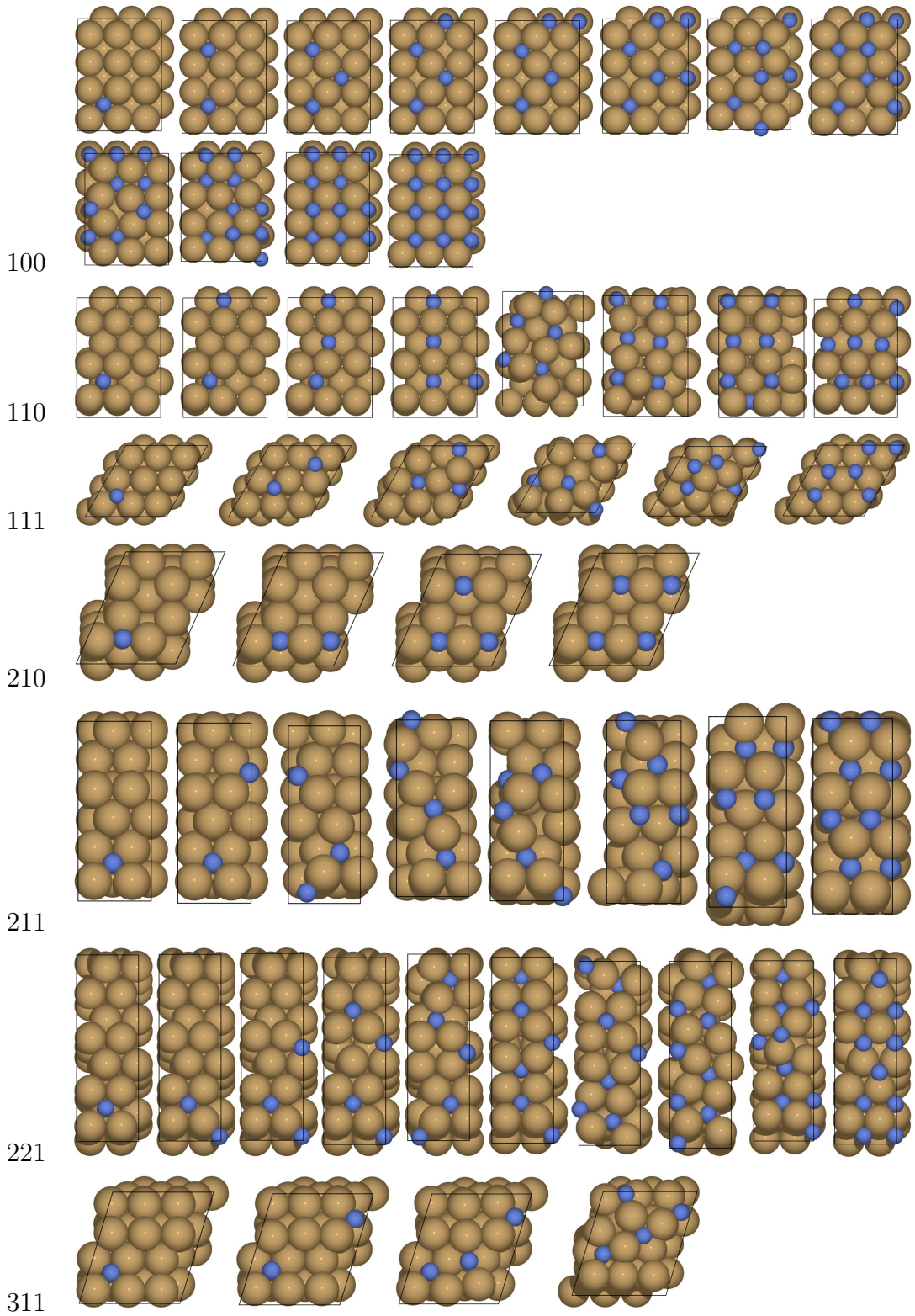


Figure A.8: Configuration search iterations for Cu/N system

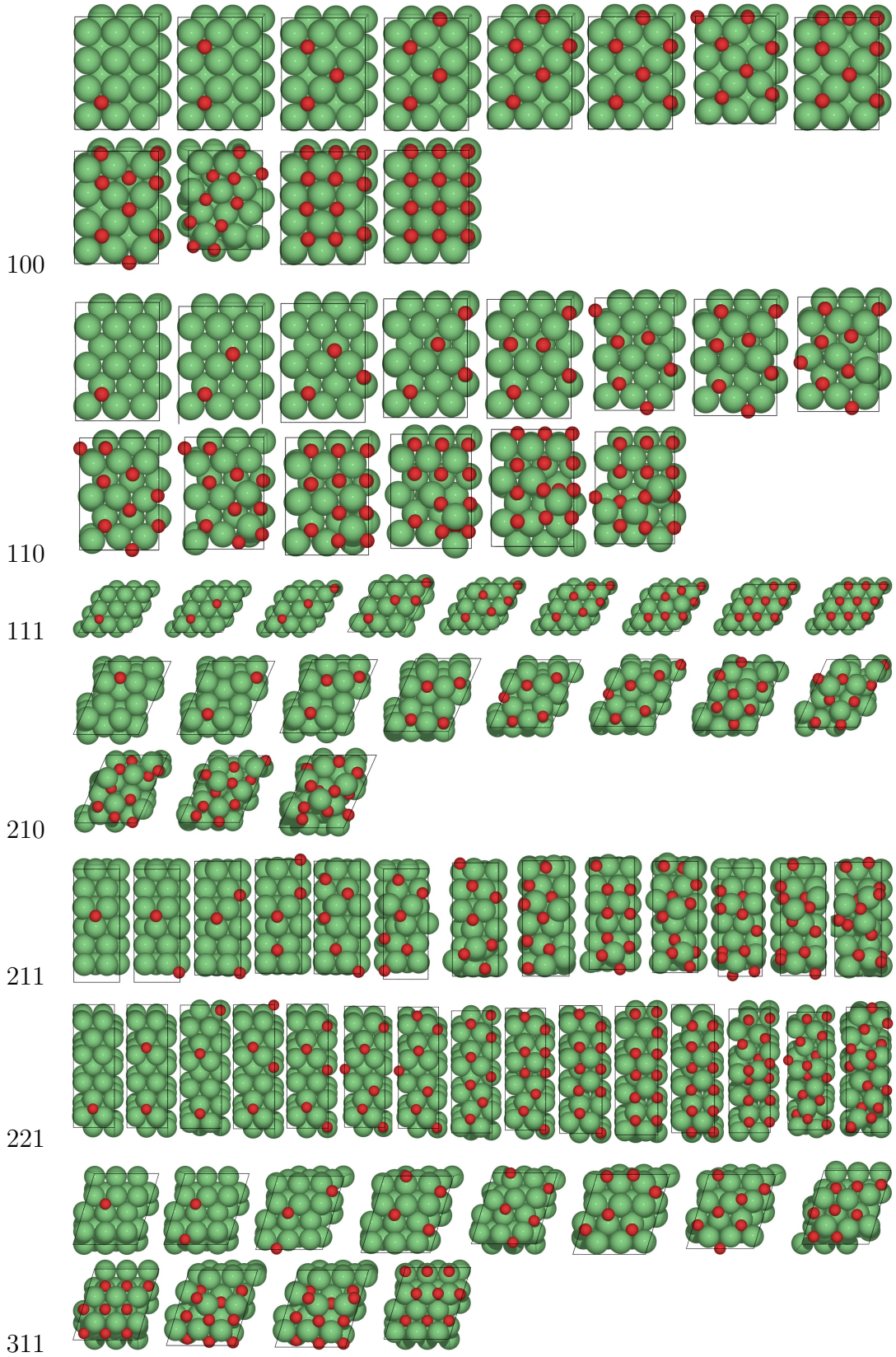


Figure A.9: Configuration search iterations for Ni/O system

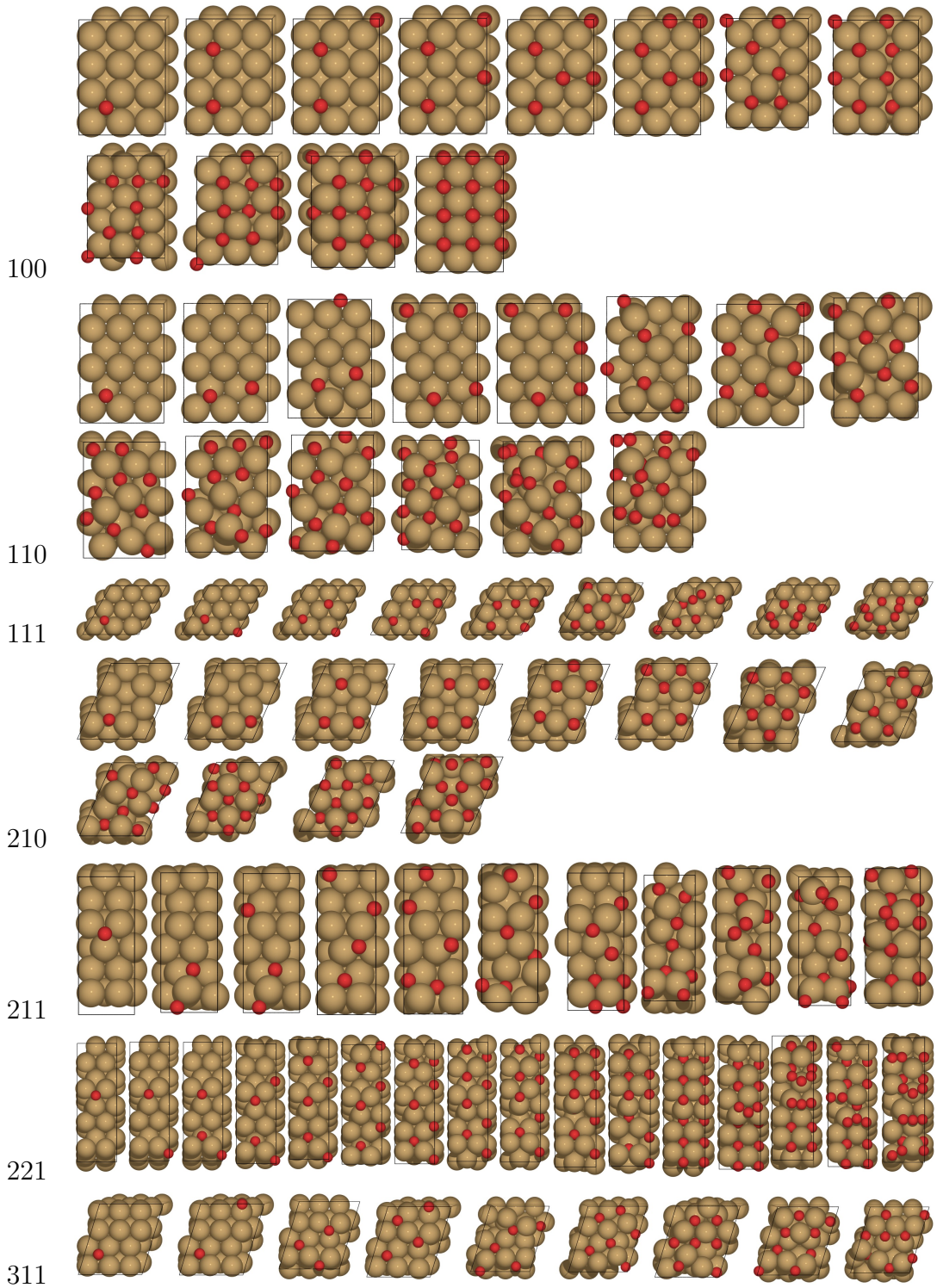


Figure A.10: Configuration search iterations for Cu/O system

A.6 Main results on Ni etching using monomer as product

We noted that the thermodynamically most stable product may not be possible if kinetic barriers are significant. Hence we also provide the figures for the case where the nickel monomers are formed as opposed to the dimer and trimer. Changing the product only affects the absolute scale of the etching energy. Figure A.11 reproduces figure 3.4 when the nickel monomer complex is produced. Neither oxygen nor nitrogen can lead to a sufficiently activated overlayer adsorption structure that is favorable to etching on any of the terminations investigated.

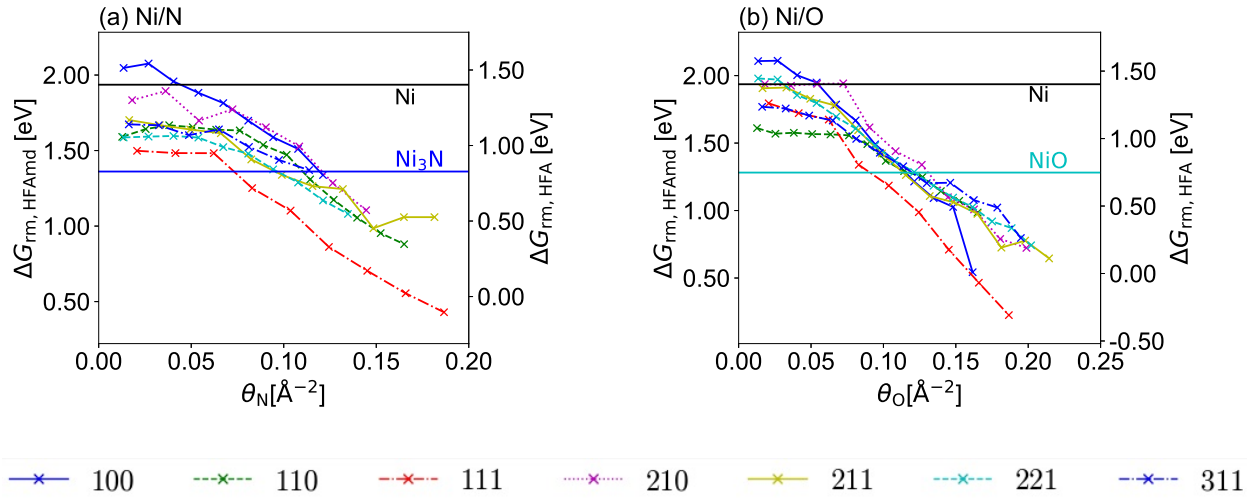


Figure A.11: Removal step energies by terminations on the Ni/N, Ni/O systems as functions of coverage using nickel formate and formamidinate monomers as products. Two axis are shown. The y-axis on the left corresponds to formamidine chemistry. The y-axis on the right corresponds to formic acid chemistry. The solid and dashed horizontal lines represents the bulk model results on the pristine and activated substrates, respectively.

Similarly, figure A.12 reproduces figure 3.5 with monomer complex as the product. Here, the dramatic effect of surface morphology is highlighted with crosses showing our previous results with metastable surface structures. It can be seen that the occupation of sub-surface sites are required for a favorable etch with the monomer complex as the product.

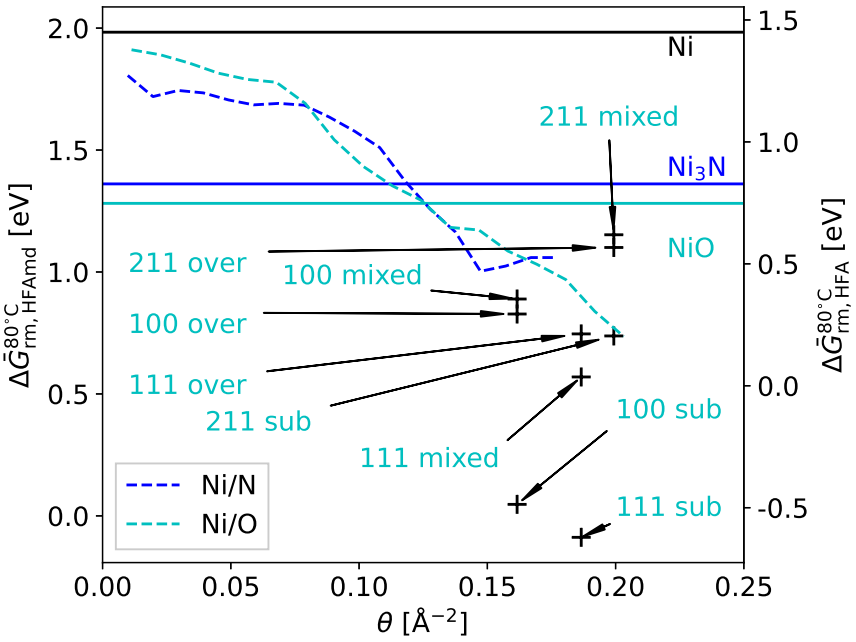


Figure A.12: Wulff construction-averaged removal step energies as functions of coverage for nickel system using nickel formate and formamidinate monomers as products.

Appendix B

Supplementary information for the neural network potential

```
1000 ###############################################################################
     # DATA SET NORMALIZATION
1002 ###############################################################################
     # This section was automatically added by nnp-norm.
1004 mean_energy  -4.0259082192812095E+00
     conv_energy   1.6694596963626187E+00
1006 conv_length   3.8772079020784176E+00
     ###############################################################################
1008
     ###############################################################################
1010 # GENERAL NNP SETTINGS
     ###############################################################################
1012 # These keywords are (almost) always required.

1014 number_of_elements              2              # Number of elements.
     elements                        O Cu           # Specification of elements.
1016 atom_energy                     O  0.0         # Free atom reference energy (S).
     atom_energy                     Cu 0.0         # Free atom reference energy (Cu).
1018 cutoff_type                     1 0.0          # Cutoff type (optional argument: shift parameter alpha)

1020 # scale_symmetry_functions                      # Scale all symmetry functions with min/max values.
     # center_symmetry_functions                    # Center all symmetry functions, i.e. subtract mean value.
1022 scale_symmetry_functions_sigma                 # Scale all symmetry functions with sigma.
     scale_min_short                 0.0            # Minimum value for scaling.
1024 scale_max_short                 1.0            # Maximum value for scaling.

1026 global_hidden_layers_short      2              # Number of hidden layers.
     global_nodes_short              30 30          # Number of nodes in each hidden layer.
1028 global_activation_short         s s l          # Activation function for each hidden layer and output layer.
     # normalize_nodes                              # Normalize input of nodes.
1030
     ###############################################################################
1032 # ADDITIONAL SETTINGS FOR DATASET TOOLS
     ###############################################################################
1034 # These keywords are used only by some tools handling data sets:
     # nnp-comp2, nnp-scaling, nnp-dataset, nnp-train.
1036
     use_short_forces                               # Use forces.
1038 random_seed                    1234567        # Random number generator seed.

1040 ###############################################################################
     # Radial symmetry function set, for elements ['O', 'Cu']
1042 ###############################################################################
```

```
# r_cutoff       = 6.0
# The following settings were used for generating sets
# of values for the radial parameters r_shift and eta:
# rule           = imbalzano2018
# mode           = center
# nb_param_pairs = 6
# Sets of values for parameters:
# r_shift_grid   = [0. 0. 0. 0. 0. 0.]
# eta_grid       = [0.0278 0.0529 0.1007 0.1916 0.3648 0.6944]

symfunction_short O  2 O  2.778E-02 0.000E+00 6.000E+00
symfunction_short O  2 O  5.288E-02 0.000E+00 6.000E+00
symfunction_short O  2 O  1.007E-01 0.000E+00 6.000E+00
symfunction_short O  2 O  1.916E-01 0.000E+00 6.000E+00
symfunction_short O  2 O  3.648E-01 0.000E+00 6.000E+00
symfunction_short O  2 O  6.944E-01 0.000E+00 6.000E+00

symfunction_short O  2 Cu 2.778E-02 0.000E+00 6.000E+00
symfunction_short O  2 Cu 5.288E-02 0.000E+00 6.000E+00
symfunction_short O  2 Cu 1.007E-01 0.000E+00 6.000E+00
symfunction_short O  2 Cu 1.916E-01 0.000E+00 6.000E+00
symfunction_short O  2 Cu 3.648E-01 0.000E+00 6.000E+00
symfunction_short O  2 Cu 6.944E-01 0.000E+00 6.000E+00

symfunction_short Cu 2 O  2.778E-02 0.000E+00 6.000E+00
symfunction_short Cu 2 O  5.288E-02 0.000E+00 6.000E+00
symfunction_short Cu 2 O  1.007E-01 0.000E+00 6.000E+00
symfunction_short Cu 2 O  1.916E-01 0.000E+00 6.000E+00
symfunction_short Cu 2 O  3.648E-01 0.000E+00 6.000E+00
symfunction_short Cu 2 O  6.944E-01 0.000E+00 6.000E+00

symfunction_short Cu 2 Cu 2.778E-02 0.000E+00 6.000E+00
symfunction_short Cu 2 Cu 5.288E-02 0.000E+00 6.000E+00
symfunction_short Cu 2 Cu 1.007E-01 0.000E+00 6.000E+00
symfunction_short Cu 2 Cu 1.916E-01 0.000E+00 6.000E+00
symfunction_short Cu 2 Cu 3.648E-01 0.000E+00 6.000E+00
symfunction_short Cu 2 Cu 6.944E-01 0.000E+00 6.000E+00

########################################################################
# Radial symmetry function set, for elements ['O', 'Cu']
########################################################################
# r_cutoff       = 6.0
# The following settings were used for generating sets
# of values for the radial parameters r_shift and eta:
# rule           = imbalzano2018
# mode           = shift
# nb_param_pairs = 4
# Sets of values for parameters:
# r_shift_grid   = [1.5    2.1213 3.     4.2426]
# eta_grid       = [2.5904 1.2952 0.6476 0.3238]

symfunction_short O  2 O  2.590E+00 1.500E+00 6.000E+00
symfunction_short O  2 O  1.295E+00 2.121E+00 6.000E+00
symfunction_short O  2 O  6.476E-01 3.000E+00 6.000E+00
symfunction_short O  2 O  3.238E-01 4.243E+00 6.000E+00

symfunction_short O  2 Cu 1.295E+00 2.121E+00 6.000E+00
symfunction_short O  2 Cu 6.476E-01 3.000E+00 6.000E+00
symfunction_short O  2 Cu 3.238E-01 4.243E+00 6.000E+00

```

```
symfunction_short Cu 2 O  2.590E+00 1.500E+00 6.000E+00
symfunction_short Cu 2 O  1.295E+00 2.121E+00 6.000E+00
symfunction_short Cu 2 O  6.476E-01 3.000E+00 6.000E+00
symfunction_short Cu 2 O  3.238E-01 4.243E+00 6.000E+00

symfunction_short Cu 2 Cu 1.295E+00 2.121E+00 6.000E+00
symfunction_short Cu 2 Cu 6.476E-01 3.000E+00 6.000E+00
symfunction_short Cu 2 Cu 3.238E-01 4.243E+00 6.000E+00

########################################################################
# Narrow angular symmetry function set, for elements ['O', 'Cu']
########################################################################
# r_cutoff       = 6.0
# The following settings were used for generating sets
# of values for the radial parameters r_shift and eta:
# rule           = imbalzano2018
# mode           = center
# nb_param_pairs = 6
# Sets of values for parameters:
# r_shift_grid   = [0. 0. 0. 0. 0. 0.]
# eta_grid       = [0.0278 0.0529 0.1007 0.1916 0.3648 0.6944]
# lambdas        = [-1.  1.]
# zetas          = [ 1.  4. 16.]

symfunction_short O  3 O  O  2.778E-02 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  2.778E-02  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  2.778E-02 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  2.778E-02  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  5.288E-02 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  5.288E-02  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  5.288E-02 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  5.288E-02  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  5.288E-02  1 1.600E+01 6.000E+00 0.000E+00
symfunction_short O  3 O  O  1.007E-01 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  1.007E-01  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  1.007E-01 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  1.007E-01  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  1.007E-01  1 1.600E+01 6.000E+00 0.000E+00
symfunction_short O  3 O  O  1.916E-01 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  1.916E-01  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  1.916E-01 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  1.916E-01  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  1.916E-01  1 1.600E+01 6.000E+00 0.000E+00
symfunction_short O  3 O  O  3.648E-01 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  3.648E-01  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  3.648E-01 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  3.648E-01  1 4.000E+00 6.000E+00 0.000E+00

symfunction_short O  3 O  O  6.944E-01 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  6.944E-01  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  6.944E-01 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  6.944E-01  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  O  6.944E-01 -1 1.600E+01 6.000E+00 0.000E+00

symfunction_short O  3 O  Cu 2.778E-02 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 2.778E-02  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 2.778E-02 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 2.778E-02  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 2.778E-02 -1 1.600E+01 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 2.778E-02  1 1.600E+01 6.000E+00 0.000E+00
```

```
symfunction_short O  3 O  Cu 5.288E-02 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 5.288E-02  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 5.288E-02 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 5.288E-02  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 5.288E-02 -1 1.600E+01 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 5.288E-02  1 1.600E+01 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 1.007E-01 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 1.007E-01  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 1.007E-01 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 1.007E-01  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 1.007E-01 -1 1.600E+01 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 1.007E-01  1 1.600E+01 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 1.916E-01 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 1.916E-01  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 1.916E-01 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 1.916E-01  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 1.916E-01  1 1.600E+01 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 3.648E-01  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 O  Cu 3.648E-01  1 4.000E+00 6.000E+00 0.000E+00

symfunction_short O  3 Cu Cu 2.778E-02 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 Cu Cu 2.778E-02  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 Cu Cu 2.778E-02 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 Cu Cu 2.778E-02  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 Cu Cu 2.778E-02  1 1.600E+01 6.000E+00 0.000E+00
symfunction_short O  3 Cu Cu 5.288E-02 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 Cu Cu 5.288E-02  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 Cu Cu 5.288E-02 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 Cu Cu 5.288E-02  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 Cu Cu 5.288E-02 -1 1.600E+01 6.000E+00 0.000E+00
symfunction_short O  3 Cu Cu 5.288E-02  1 1.600E+01 6.000E+00 0.000E+00
symfunction_short O  3 Cu Cu 1.007E-01 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 Cu Cu 1.007E-01  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 Cu Cu 1.007E-01 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 Cu Cu 1.007E-01  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 Cu Cu 1.007E-01 -1 1.600E+01 6.000E+00 0.000E+00
symfunction_short O  3 Cu Cu 1.007E-01  1 1.600E+01 6.000E+00 0.000E+00
symfunction_short O  3 Cu Cu 1.916E-01 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 Cu Cu 1.916E-01  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 Cu Cu 1.916E-01 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 Cu Cu 1.916E-01  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 Cu Cu 3.648E-01 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short O  3 Cu Cu 3.648E-01  1 1.000E+00 6.000E+00 0.000E+00

symfunction_short Cu 3 O  O  2.778E-02 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  2.778E-02  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  2.778E-02 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  2.778E-02  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  2.778E-02 -1 1.600E+01 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  5.288E-02 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  5.288E-02  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  5.288E-02 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  5.288E-02  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  5.288E-02 -1 1.600E+01 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  5.288E-02  1 1.600E+01 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  1.007E-01 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  1.007E-01  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  1.007E-01 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  1.007E-01  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  1.007E-01 -1 1.600E+01 6.000E+00 0.000E+00
```

```
symfunction_short Cu 3 O  O  1.916E-01 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  1.916E-01  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  1.916E-01 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  1.916E-01  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  1.916E-01 -1 1.600E+01 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  3.648E-01 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  3.648E-01  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  3.648E-01 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  3.648E-01  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  6.944E-01 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  6.944E-01  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  O  6.944E-01  1 4.000E+00 6.000E+00 0.000E+00

symfunction_short Cu 3 O  Cu 2.778E-02 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 2.778E-02  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 2.778E-02 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 2.778E-02  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 2.778E-02 -1 1.600E+01 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 2.778E-02  1 1.600E+01 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 5.288E-02 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 5.288E-02  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 5.288E-02 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 5.288E-02  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 5.288E-02 -1 1.600E+01 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 5.288E-02  1 1.600E+01 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 1.007E-01 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 1.007E-01  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 1.007E-01 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 1.007E-01  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 1.007E-01 -1 1.600E+01 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 1.007E-01  1 1.600E+01 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 1.916E-01 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 1.916E-01  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 1.916E-01  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 1.916E-01 -1 1.600E+01 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 1.916E-01  1 1.600E+01 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 3.648E-01  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 3.648E-01  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 6.944E-01  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 O  Cu 6.944E-01  1 4.000E+00 6.000E+00 0.000E+00

symfunction_short Cu 3 Cu Cu 2.778E-02 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 Cu Cu 2.778E-02  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 Cu Cu 2.778E-02 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 Cu Cu 2.778E-02  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 Cu Cu 2.778E-02 -1 1.600E+01 6.000E+00 0.000E+00
symfunction_short Cu 3 Cu Cu 2.778E-02  1 1.600E+01 6.000E+00 0.000E+00
symfunction_short Cu 3 Cu Cu 5.288E-02 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 Cu Cu 5.288E-02  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 Cu Cu 5.288E-02 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 Cu Cu 5.288E-02  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 Cu Cu 5.288E-02 -1 1.600E+01 6.000E+00 0.000E+00
symfunction_short Cu 3 Cu Cu 5.288E-02  1 1.600E+01 6.000E+00 0.000E+00
symfunction_short Cu 3 Cu Cu 1.007E-01 -1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 Cu Cu 1.007E-01  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 Cu Cu 1.007E-01 -1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 Cu Cu 1.007E-01  1 4.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 Cu Cu 1.007E-01 -1 1.600E+01 6.000E+00 0.000E+00
symfunction_short Cu 3 Cu Cu 1.007E-01  1 1.600E+01 6.000E+00 0.000E+00
symfunction_short Cu 3 Cu Cu 1.916E-01 -1 1.000E+00 6.000E+00 0.000E+00
```

```
symfunction_short Cu 3 Cu Cu 1.916E-01  1 1.000E+00 6.000E+00 0.000E+00
symfunction_short Cu 3 Cu Cu 1.916E-01  1 4.000E+00 6.000E+00 0.000E+00
```

Listing B.1: neural network potential architecture

Appendix C

Supplementary information for the molecular dynamics simulation

C.1 Post-processing and analysis of MD trajectories

C.1.1 Calculation of oxide film thickness

The film thicknesses calculation involves the following steps:

1. locating the top surface by identifying the interfacial atoms. This is done using the pytim module, with the method described in Ref.[78].

2. The bottom surface is located in a similar fashion. However, the method cannot be used to determine interfacial atoms where the density stays more or less constant but the composition changes abruptly. Therefore, all the copper atoms were removed prior to detecting the interfacial oxygen atoms.

3. The average z-coordinate of the atoms in the top and bottom surface are calculated, and the difference $\delta z = \bar{z}_{\text{top}} - \bar{z}_{\text{bottom}}$ is reported as the film thickness. In very thin films the top and bottom surface atoms overlap. In such cases the thicknesses is directly set to 0.

C.1.2 Characterization of oxide composition

To determine the chemical composition of the oxide film, the local environments of each atom were analyzed using bond orientational order parameters.[12] The Voronoi tessellation

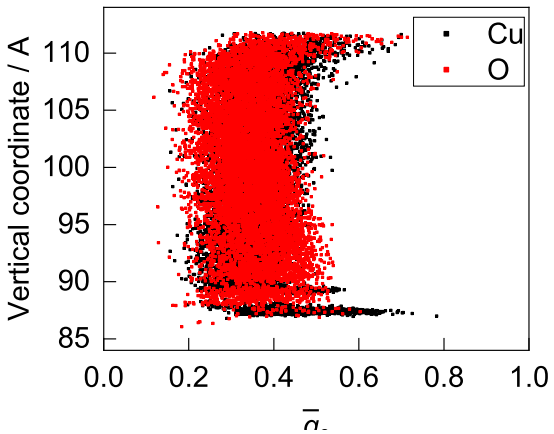


Figure C.1: q_6 parameters showing overlap of oxygen and copper atoms

averaged version, shown below, is used.[97]

$$q_l'(a) = \sqrt{\frac{4\pi}{2l+1}\sum_{m=-l}^{l}|\frac{A(f)}{A}\sum_{b\in \mathrm{NN}(a)}Y_{lm}(\theta_{ab},\psi_{ab})|^2}$$

, where Y_{lm} is a spherical harmonic function of degree l and order m, θ_{ab} and ψ_{ab} are polar and azimuthal angles of the distance vector between neighboring atoms in spherical coordinates, $A(f)$ and A are the separating surface in the Voronoi tessellation between neighboring atoms and the total area of the Voronoi cell of the center atom, respectively.

C.1.3 Tracing the range of thermal agitations

Starting with the detection of the collision event detailed above, the atoms near the impact site are monitored for their velocities. The atom with velocities exceeding the cutoff velocity is added to the tracked hot atoms. The algorithm is repeated recursively in the list of "hot" atoms for each time step, therefore accounting for all the atoms affected by the thermal agitation. The cutoff velocity is a parameter specified in terms of the probability of particles with such velocities being present in the integrated Maxwell-Boltzmann distribution, which the velocity distributions in a amorphous solid are assumed to have.[141] In this work, this cutoff probability is 1×10^{-5}

C.2 Comparison of simulated and experimental impact rates

Given the pressure of the plasma processing step, the impact frequency can be calculated using the kinetic theory of gases:

$$J_{\text{collision}} = \frac{P}{\sqrt{2\pi m k_B T}}$$

At a pressure of 30 mTorr, $J_{\text{collision}} = 4.378 \times 10^{-10} \frac{1}{\text{nm}^2\text{fs}}$. The slab used during production MD had 26 nm^2 area, corresponding to an impact frequency of $1.14 \times 10^{-8}\,\text{fs}^{-}1$. Therefore, to capture even one impact event at laboratory time scale, 87.8 ns is needed, clearly beyond current capabilities. At best, the MD simulations of plasma processing may be related to laboratory times via fluences of the ions. For instance, for 353K/20eV/20, 23314100 MD steps were performed. For every 10000 steps, 21 oxygen atoms (both ions and neutrals) were deposited, leading to a total of 46628 ions. Since the impingement frequency is calculated to be 87.8 ns per ion/neutral, equating the fluences would correspond to 4.09×10^{-3} s of laboratory time, which is still much less than the real exposure time ($\sim$ 1 min). This four order of magnitude difference in between can only be bridged by accelerated MD or hybrid MD/MC(Monte Carlo) methods, and is the most significant challenge moving forward.

C.3 Additional results on the 111 surface

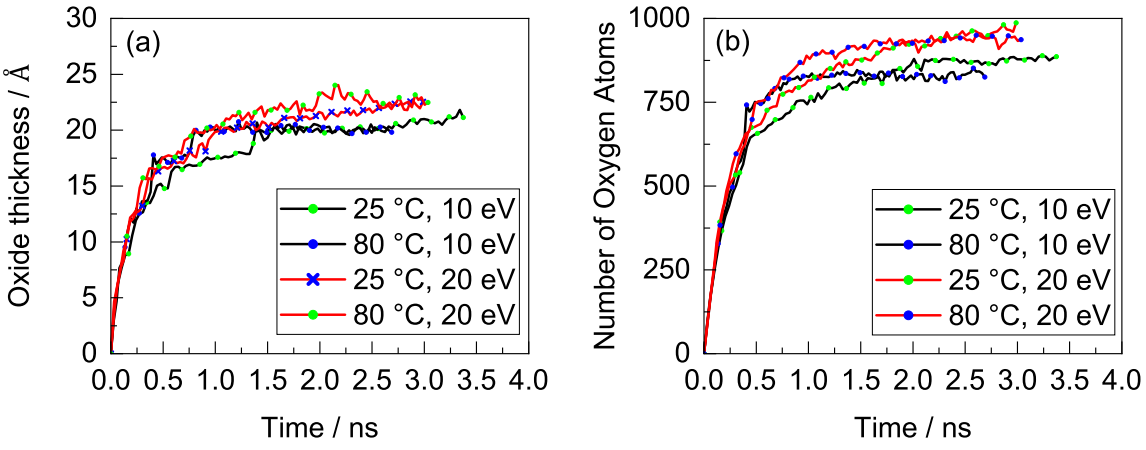


Figure C.2: Thickness and number of oxygen in the growing oxide on Cu (111) surface.

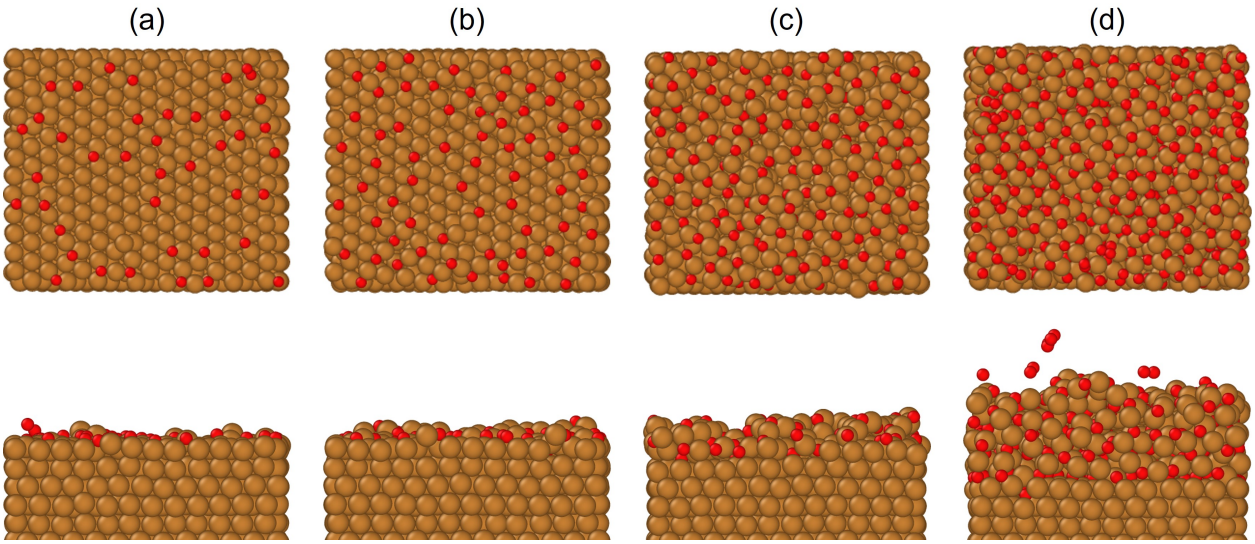


Figure C.3: Oxidation trajectory on the Cu (111) surface

Bibliography

(1) Boyle, B. J.; King, E. G.; Conway, K. C. *Journal of the American Chemical Society* **1954**, *76*, 3835–3837.

(2) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864–B871.

(3) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133–A1138.

(4) Edwards, D.; Richards, R. *Inorganic and Nuclear Chemistry Letters* **1972**, *8*, 783–792.

(5) Ogura, T.; Fernando, Q. *Inorganic Chemistry* **1973**, *12*, 2611–2615.

(6) Ogura, T.; Mounts, R. D.; Fernando, Q. *Journal of the American Chemical Society* **1973**, *95*, 949–951.

(7) Andersen, O. K. *Physical Review B* **1975**, *12*, Publisher: American Physical Society, 3060–3083.

(8) Chapman, B., *Glow Discharge Processes: Sputtering and Plasma Etching*; Wiley: 1980.

(9) Narusawa, T.; Gibson, W.; Törnqvist, E. *Surface Science* **1982**, *114*, 331–348.

(10) Baró, A.; Ollé, L. *Surface Science* **1983**, *126*, 170–176.

(11) Narula, M. L.; Tare, V. B.; Worrell, W. L. *Metallurgical Transactions B* **1983**, *14*, 673–677.

(12) Steinhardt, P. J.; Nelson, D. R.; Ronchetti, M. *Physical Review B* **1983**, *28*, Publisher: American Physical Society, 784–805.

(13) Daw, M. S.; Baskes, M. I. *Physical Review B* **1984**, *29*, 6443–6453.

(14) Matsunami, N.; Yamamura, Y.; Itikawa, Y.; Itoh, N.; Kazumata, Y.; Miyagawa, S.; Morita, K.; Shimizu, R.; Tawara, H. *Atomic Data and Nuclear Data Tables* **1984**, *31*, 1–80.

(15) Stillinger, F. H.; Weber, T. A. *Physical Review B* **1985**, *31*, 5262–5271.

(16) Baroni, S.; Giannozzi, P.; Testa, A. *Phys. Rev. Lett.* **1987**, *58*, 1861–1864.

(17) Tersoff, J. *Physical Review B* **1988**, *37*, 6991–7000.

(18) Brenner, D. W. *Physical Review B* **1990**, *42*, 9458–9471.

(19) Vanderbilt, D. *Physical Review B* **1990**, *41*, Publisher: American Physical Society, 7892–7895.

(20) Fichthorn, K. A.; Weinberg, W. H. *The Journal of Chemical Physics* **1991**, *95*, Publisher: American Institute of Physics, 1090–1096.

(21) Rappe, A. K.; Goddard, W. A. *The Journal of Physical Chemistry* **1991**, *95*, 3358–3363.

(22) Baskes, M. I. *Physical Review B* **1992**, *46*, 2727–2742.

(23) Edwards, H. G. M.; Knowles, A. *Journal of Molecular Structure* **1992**, *268*, 13–22.

(24) Schäfer, A.; Horn, H.; Ahlrichs, R. *The Journal of Chemical Physics* **1992**, *97*, 2571–2577.

(25) Elder, S. H.; DiSalvo, F. J.; Topor, L.; Navrotsky, A. *Chemistry of Materials* **1993**, *5*, 1545–1553.

(26) Feil, H.; Dieleman, J.; Garrison, B. J. *Journal of Applied Physics* **1993**, *74*, Publisher: American Institute of Physics, 1303–1309.

(27) Blöchl, P. E. *Phys. Rev. B* **1994**, *50*, 17953–17979.

(28) Kresse, G.; Hafner, J. *Journal of Physics: Condensed Matter* **1994**, *6*, 8245.

(29) Lin, J.; Neoh, K. G.; Teo, W. k. *J. Chem. Soc., Faraday Trans.* **1994**, *90*, 355–362.

(30) Parr, R. G.; Weitao, Y., *Density-Functional Theory of Atoms and Molecules*, Google-Books-ID: mGOpScSIwU4C; Oxford University Press: 1994.

(31) Athavale, S. D.; Economou, D. J. *Journal of Vacuum Science & Technology A* **1995**, *13*, 966–971.

(32) Athavale, S. D.; Economou, D. J. *Journal of Vacuum Science & Technology A* **1995**, *13*, 966–971.

(33) Porezag, D.; Frauenheim, T.; Köhler, T.; Seifert, G.; Kaschner, R. *Physical Review B* **1995**, *51*, 12947–12957.

(34) Kresse, G.; Furthmüller, J. *Computational Materials Science* **1996**, *6*, 15–50.

(35) Kresse, G.; Furthmüller, J. *Phys. Rev. B* **1996**, *54*, 11169–11186.

(36) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.

(37) Szabo, A.; Ostlund, N. S., *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*, Google-Books-ID: 6mV9gYzEkgIC; Courier Corporation: 1996.

(38) Chang, J. P.; Arnold, J. C.; Zau, G. C. H.; Shin, H.-S.; Sawin, H. H. *Journal of Vacuum Science & Technology A* **1997**, *15*, 1853–1863.

(39) Edelstein, D. et al. In *International Electron Devices Meeting. IEDM Technical Digest*; ISSN: 0163-1918, 1997, pp 773–776.

(40) Hoekstra, R. J.; Grapperhaus, M. J.; Kushner, M. J. *Journal of Vacuum Science & Technology A* **1997**, *15*, Publisher: American Vacuum Society, 1913–1921.

(41) Voter, A. F. *The Journal of Chemical Physics* **1997**, *106*, Publisher: American Institute of Physics, 4665–4677.

(42) Voter, A. F. *Physical Review Letters* **1997**, *78*, Publisher: American Physical Society, 3908–3911.

(43) Brown, W. A.; Kose, R.; King, D. A. *Chemical Reviews* **1998**, *98*, 797–832.

(44) Bussery-Honvault, B.; Veyret, V. *The Journal of Chemical Physics* **1998**, *108*, 3243–3248.

(45) Yang, J. C.; Kolasa, B.; Gibson, J. M.; Yeadon, M. *Applied Physics Letters* **1998**, *73*, 2841–2843.

(46) Aquilanti, V.; Ascenzi, D.; Bartolomei, M.; Cappelletti, D.; Cavalli, S.; de Castro Vítores, M.; Pirani, F. *Physical Review Letters* **1999**, *82*, 69–72.

(47) Driver, S. M.; Woodruff, D. P. *Surface Science* **1999**, *442*, 1–8.

(48) Kresse, G.; Joubert, D. *Phys. Rev. B* **1999**, *59*, 1758–1775.

(49) Wang, Y.; Olthoff, J. K. *Journal of Applied Physics* **1999**, *85*, 6358–6365.

(50) Economou, D. J. *Thin Solid Films* **2000**, *365*, 348–367.

(51) Rosenberg, R.; Edelstein, D. C.; Hu, C.-K.; Rodbell, K. P. *Annual Review of Materials Science* **2000**, *30*, _eprint: https://doi.org/10.1146/annurev.matsci.30.1.229, 229–262.

(52) Van Duin, A. C. T.; Dasgupta, S.; Lorant, F.; Goddard, W. A. *The Journal of Physical Chemistry A* **2001**, *105*, 9396–9409.

(53) Yagi-Watanabe, K.; Ikeda, Y.; Ishii, Y.; Inokuchi, T.; Fukutani, H. *Surface Science* **2001**, *482-485*, 128–133.

(54) Kulkarni, N. S.; DeHoff, R. T. *Journal of The Electrochemical Society* **2002**, *149*, Publisher: IOP Publishing, G620.

(55) Chen, F.; Chang, J., *Lecture Notes on Principles of Plasma Processing*; Springer US: 2003.

(56) Hernández-Lamoneda, R.; Hernández, M.; Campos-Martínez, J. *Chemical Physics Letters* **2003**, *368*, 709–716.

(57) Jansen, A. P. J. *arXiv:cond-mat/0303028* **2003**, arXiv: cond-mat/0303028.

(58) Lim, B. S.; Rahtu, A.; Gordon, R. G. *Nature Materials* **2003**, *2*, 749–754.

(59) Schweitzer, C.; Schmidt, R. *Chemical Reviews* **2003**, *103*, 1685–1758.

(60) Elliott, S. D.; Greer, J. C. *Journal of Materials Chemistry* **2004**, *14*, Publisher: The Royal Society of Chemistry, 3246–3250.

(61) Vempaire, D.; Miraglia, S.; Sulpice, A.; Ortega, L.; Hlil, E.; Fruchart, D.; Pelletier, J. *Journal of Magnetism and Magnetic Materials* **2004**, *272-276*, Proceedings of the International Conference on Magnetism (ICM 2003), E843–E844.

(62) Cramer, C., *Essentials of Computational Chemistry: Theories and Models*; Wiley: 2005.

(63) Kodera, M.; Uekusa, S.-i.; Nagano, H.; Tokushige, K.; Shima, S.; Fukunaga, A.; Mochizuki, Y.; Fukuda, A.; Hiyama, H.; Tsujimura, M.; Nagai, H.; Maekawa, K. *Journal of The Electrochemical Society* **2005**, *152*, Publisher: IOP Publishing, G506.

(64) Li, Z.; Barry, S. T.; Gordon, R. G. *Inorganic Chemistry* **2005**, *44*, 1728–1735.

(65) Nijhuis, C. A.; Jellema, E.; Sciarone, T. J. J.; Meetsma, A.; Budzelaar, P. H. M.; Hessen, B. *European Journal of Inorganic Chemistry* **2005**, *2005*, 2089–2099.

(66) Soon, A.; Todorova, M.; Delley, B.; Stampfl, C. *Phys. Rev. B* **2006**, *73*, 165424.

(67) Zenmyo, K.; Kubo, H.; Tokita, M.; Yamagata, K. *Journal of the Physical Society of Japan* **2006**, *75*, 104704.

(68) Behler, J.; Parrinello, M. *Physical Review Letters* **2007**, *98*, 146401.

(69) Rauf, S.; Sparks, T.; Ventzek, P. L. G.; Smirnov, V. V.; Stengach, A. V.; Gaynullin, K. G.; Pavlovsky, V. A. *Journal of Applied Physics* **2007**, *101*, 033308.

(70) Rauf, S.; Sparks, T.; Ventzek, P. L. G.; Smirnov, V. V.; Stengach, A. V.; Gaynullin, K. G.; Pavlovsky, V. A. *Journal of Applied Physics* **2007**, *101*, 033308.

(71) Seifert, G. *The Journal of Physical Chemistry A* **2007**, *111*, 5609–5613.

(72) *Radiation effects in solids: proceedings of the NATO Advanced Study Institute on Radiation Effects in Solids: Erice, Sicily, Italy, 17-29 July 2004*; Sickafus, K. E., Kotomin, E. A., Uberuaga, B. P., Eds.; NATO science series. II, Mathematics, physics and chemistry v. 235, Meeting Name: NATO Advanced Study Institute on Radiation Effects in Solids OCLC: ocm74650679; Springer: Dordrecht, Netherlands, 2007.

(73) Soon, A.; Wong, L.; Lee, M.; Todorova, M.; Delley, B.; Stampfl, C. *Surface Science* **2007**, *601*, 4775–4785.

(74) Wu, J.; Li, J.; Zhou, C.; Lei, X.; Gaffney, T.; Norman, J. A. T.; Li, Z.; Gordon, R.; Cheng, H. *Organometallics* **2007**, *26*, 2803–2805.

(75) *EUV Lithography*; Bakshi, V., Ed.; SPIE: 1000 20th Street, Bellingham, WA 98227-0010 USA, 2008.

(76) Campos-Martinez, J.; I. Hernandez, M.; Bartolomei, M.; Carmona-Novillo, E.; Hernandez-Lamoneda, R.; Dayou, F. In *Frontiers in quantum systems in chemistry and physics*; Progress in theorical chemistry and physics 18; Springer: Dordrecht, 2008.

(77) Juaristi, J. I.; Alducin, M.; Muiño, R. D.; Busnengo, H. F.; Salin, A. *Physical Review Letters* **2008**, *100*, 116102.

(78) Pártay, L. B.; Hantal, G.; Jedlovszky, P.; Vincze, Á.; Horvai, G. *Journal of Computational Chemistry* **2008**, *29*, 945–956.

(79) Soon, A.; Wong, L.; Delley, B.; Stampfl, C. *Physical Review B* **2008**, *77*, Publisher: American Physical Society, 125423.

(80) Graves, D. B.; Brault, P. *Journal of Physics D: Applied Physics* **2009**, *42*, 194011.

(81) Li, J.; Wu, J.; Zhou, C.; Han, B.; Lei, X.; Gordon, R.; Cheng, H. *International Journal of Quantum Chemistry* **2009**, *109*, 756–763.

(82) Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. *Physical Review Letters* **2010**, *104*, 136403.

(83) Duan, X.; Warschkow, O.; Soon, A.; Delley, B.; Stampfl, C. *Phys. Rev. B* **2010**, *81*, 075430.

(84) Gou, F.; Neyts, E.; Eckert, M.; Tinck, S.; Bogaerts, A. *Journal of Applied Physics* **2010**, *107*, 113305.

(85) Gou, F.; Neyts, E.; Eckert, M.; Tinck, S.; Bogaerts, A. *Journal of Applied Physics* **2010**, *107*, 113305.

(86) Li, Z.; Gordon, R. G.; Pallem, V.; Li, H.; Shenai, D. V. *Chemistry of Materials* **2010**, *22*, 3060–3066.

(87) Shan, T.-R.; Devine, B. D.; Kemper, T. W.; Sinnott, S. B.; Phillpot, S. R. *Physical Review B* **2010**, *81*, 125328.

(88) Van Duin, A. C. T.; Bryantsev, V. S.; Diallo, M. S.; Goddard, W. A.; Rahaman, O.; Doren, D. J.; Raymand, D.; Hermansson, K. *The Journal of Physical Chemistry A* **2010**, *114*, Publisher: American Chemical Society, 9507–9514.

(89) Devine, B.; Shan, T.-R.; Cheng, Y.-T.; McGaughey, A. J. H.; Lee, M.; Phillpot, S. R.; Sinnott, S. B. *Physical Review B* **2011**, *84*, Publisher: American Physical Society, 125308.

(90) Jeon, B.; Sankaranarayanan, S. K.; van Duin, A. C.; Ramanathan, S. *Philosophical Magazine* **2011**, *91*, 4073–4088.

(91) Moroz, P. *ECS Transactions* **2011**, *35*, 25–31.

(92) Steinmann, S. N.; Corminboeuf, C. *The Journal of Chemical Physics* **2011**, *134*, 044117.

(93) Steinmann, S. N.; Corminboeuf, C. *Journal of Chemical Theory and Computation* **2011**, *7*, PMID: 26598255, 3567–3577.

(94) Aktulga, H. M.; Fogarty, J. C.; Pandit, S. A.; Grama, A. Y. *Parallel Computing* **2012**, *38*, 245–259.

(95) Zhou, G.; Luo, L.; Li, L.; Ciston, J.; Stach, E. A.; Yang, J. C. *Physical Review Letters* **2012**, *109*, 235502.

(96) Liang, T.; Shan, T.-R.; Cheng, Y.-T.; Devine, B. D.; Noordhoek, M.; Li, Y.; Lu, Z.; Phillpot, S. R.; Sinnott, S. B. *Materials Science and Engineering: R: Reports* **2013**, *74*, 255–279.

(97) Mickel, W.; Kapfer, S. C.; Schröder-Turk, G. E.; Mecke, K. *The Journal of Chemical Physics* **2013**, *138*, 044501.

(98) Obata, M.; Nakamura, M.; Hamada, I.; Oda, T. *Journal of the Physical Society of Japan* **2013**, *82*, 093701.

(99) Kühne, T. D. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2014**, *4*, 391–406.

(100) Magnuson, H.; Frisk, K. *Self-diffusion and impurity diffusion of hydrogen, oxygen, sulphur and phosphorus in copper*; tech. rep. TR-13-24, Swedish Nuclear Fuel and Waste Management Company; Svensk Kärnbränslehantering Aktiebolag, 2014, pp 14–16.

(101) Moroz, P.; Moroz, D. J. *Journal of Physics: Conference Series* **2014**, *550*, 012030.

(102) Pedersen, H.; Elliott, S. D. *Theoretical Chemistry Accounts* **2014**, *133*, 1476.

(103) Bal, K. M.; Neyts, E. C. *Journal of Chemical Theory and Computation* **2015**, *11*, 4545–4554.

(104) Carver, C. T.; Plombon, J. J.; Romero, P. E.; Suri, S.; Tronic, T. A.; Turkot, R. B. *ECS Journal of Solid State Science and Technology* **2015**, *4*, N5005–N5009.

(105) Carver, C. T.; Plombon, J. J.; Romero, P. E.; Suri, S.; Tronic, T. A.; Turkot, R. B. *ECS Journal of Solid State Science and Technology* **2015**, *4*, N5005–N5009.

(106) Gattinoni, C.; Michaelides, A. *Surface Science Reports* **2015**, *70*, 424–447.

(107) Gautier, S.; Steinmann, S. N.; Michel, C.; Fleurat-Lessard, P.; Sautet, P. *Physical Chemistry Chemical Physics* **2015**, *17*, Publisher: The Royal Society of Chemistry, 28921–28930.

(108) Kanarik, K. J.; Lill, T.; Hudson, E. A.; Sriraman, S.; Tan, S.; Marks, J.; Vahedi, V.; Gottscho, R. A. *Journal of Vacuum Science & Technology A* **2015**, *33*, 020802.

(109) Lee, Y.; George, S. M. *ACS Nano* **2015**, *9*, 2061–2070.

(110) Tan, S.; Yang, W.; Kanarik, K. J.; Lill, T.; Vahedi, V.; Marks, J.; Gottscho, R. A. *ECS Journal of Solid State Science and Technology* **2015**, *4*, N5010–N5012.

(111) Zagidullin, M. V.; Pershin, A. A.; Azyazov, V. N.; Mebel, A. M. *The Journal of Chemical Physics* **2015**, *143*, 244315.

(112) Zhu, Q.; Saidi, W. A.; Yang, J. C. *The Journal of Physical Chemistry C* **2015**, *119*, 251–261.

(113) Frisch, M. J. et al. Gaussian˜16 Revision C.01, Gaussian Inc. Wallingford CT, 2016.

(114) Hay, D.; Bagge, P.; Khaw, I.; Sun, L.; Wood, O.; Chen, Y.; Kim, R.-h.; Qi, Z. J.; Shi, Z. *Optics Letters* **2016**, *41*, Publisher: Optical Society of America, 3791–3794.

(115) Haynes, W. M. In 97th ed.; CRC Press: 2016, pp 5–11, 5–29, 5–30.

(116) Hellman, A. *Chemistry of Materials* **2016**.

(117) Johnson, N. R.; Sun, H.; Sharma, K.; George, S. M. *Journal of Vacuum Science & Technology A* **2016**, *34*, Publisher: American Vacuum Society, 050603.

(118) Karton, A. *WIREs Computational Molecular Science* **2016**, *6*, _eprint: https://onlinelibrary.wiley.co 292–310.

(119) Lee, Y.; DuMont, J. W.; George, S. M. *Chemistry of Materials* **2016**, *28*, Publisher: American Chemical Society, 2994–3003.

(120) Senftle, T. P.; Hong, S.; Islam, M. M.; Kylasa, S. B.; Zheng, Y.; Shin, Y. K.; Junkermeier, C.; Engel-Herbert, R.; Janik, M. J.; Aktulga, H. M.; Verstraelen, T.; Grama, A.; van Duin, A. C. T. *npj Computational Materials* **2016**, *2*, 15011.

(121) Silvestrelli, P. L.; Ambrosetti, A. *Physical Review B* **2016**, *94*, Publisher: American Physical Society, 045124.

(122) Tiwary, P.; van de Walle, A. In *Multiscale Materials Modeling for Nanomechanics*, Weinberger, C. R., Tucker, G. J., Eds.; Springer Series in Materials Science; Springer International Publishing: Cham, 2016, pp 195–221.

(123) Zhai, H.; Alexandrova, A. N. *Journal of Chemical Theory and Computation* **2016**, *12*, 6213–6226.

(124) Zhu, Q.; Zou, L.; Zhou, G.; Saidi, W. A.; Yang, J. C. *Surface Science* **2016**, *652*, 98–113.

(125) Adamovich, I. et al. *Journal of Physics D: Applied Physics* **2017**, *50*, 323001.

(126) Altieri, N. D.; Chen, J. K.-C.; Minardi, L.; Chang, J. P. *Journal of Vacuum Science & Technology A* **2017**, *35*, 05C203.

(127) Aussems, D. U. B.; Bal, K. M.; Morgan, T. W.; Sanden, M. C. M. v. d.; Neyts, E. C. *Chemical Science* **2017**, *8*, Publisher: The Royal Society of Chemistry, 7160–7168.

(128) Chen, J. K.-C.; Altieri, N. D.; Kim, T.; Chen, E.; Lill, T.; Shen, M.; Chang, J. P. *Journal of Vacuum Science & Technology A* **2017**, *35*, 05C305.

(129) Chen, J. K.-C.; Altieri, N. D.; Kim, T.; Lill, T.; Shen, M.; Chang, J. P. *Journal of Vacuum Science & Technology A* **2017**, *35*, 05C304.

(130) Edelstein, D. C. In *2017 IEEE International Electron Devices Meeting (IEDM)*; ISSN: 2156-017X, 2017, pp 14.1.1–14.1.4.

(131) Hu, X.; Schuster, J.; Schulz, S. E. *The Journal of Physical Chemistry C* **2017**, *121*, 28077–28089.

(132) Larsen, A. H. et al. *Journal of Physics: Condensed Matter* **2017**, *29*, Publisher: IOP Publishing, 273002.

(133) Neyts, E. C.; Brault, P. *Plasma Processes and Polymers* **2017**, *14*, 1600145.

(134) Scholze, F.; Laubis, C.; Luong, K. V.; Philipsen, V. In *33rd European Mask and Lithography Conference*, International Society for Optics and Photonics: 2017; Vol. 10446, p 1044609.

(135) Choudhary, S.; Sarma, J. V. N.; Pande, S.; Ababou-Girard, S.; Turban, P.; Lepine, B.; Gangopadhyay, S. *AIP Advances* **2018**, *8*, 055114.

(136) Deringer, V. L.; Bernstein, N.; Bartók, A. P.; Cliffe, M. J.; Kerber, R. N.; Marbella, L. E.; Grey, C. P.; Elliott, S. R.; Csányi, G. *The Journal of Physical Chemistry Letters* **2018**, *9*, Publisher: American Chemical Society, 2879–2885.

(137) Fang, C.; Cao, Y.; Wu, D.; Li, A. *Progress in Natural Science: Materials International* **2018**, *28*, 667–675.

(138) Imbalzano, G.; Anelli, A.; Giofré, D.; Klees, S.; Behler, J.; Ceriotti, M. *The Journal of Chemical Physics* **2018**, *148*, 241730.

(139) Kondati Natarajan, S.; Elliott, S. D. *Chemistry of Materials* **2018**, *30*, 5912–5922.

(140) Li, C.; Zhang, P.; Wang, J.; Boscoboinik, J. A.; Zhou, G. *The Journal of Physical Chemistry C* **2018**, *122*, 8254–8261.

(141) Mohazzabi, P.; Shankar, S. P. *Journal of Applied Mathematics and Physics* **2018**, *6*, 602–612.

(142) Yeoh, A. et al. In *2018 IEEE International Interconnect Technology Conference (IITC)*; ISSN: 2380-6338, 2018, pp 144–147.

(143) Borlido, P.; Aull, T.; Huran, A. W.; Tran, F.; Marques, M. A. L.; Botti, S. *Journal of Chemical Theory and Computation* **2019**, *15*, Publisher: American Chemical Society, 5069–5079.

(144) Konh, M.; He, C.; Lin, X.; Guo, X.; Pallem, V.; Opila, R. L.; Teplyakov, A. V.; Wang, Z.; Yuan, B. *Journal of Vacuum Science & Technology A* **2019**, *37*, 021004.

(145) Ohmi, H.; Sato, J.; Shirasu, Y.; Hirano, T.; Kakiuchi, H.; Yasutake, K. *ACS Omega* **2019**, *4*, Publisher: American Chemical Society, 4360–4366.

(146) Singraber, A.; Behler, J.; Dellago, C. *Journal of Chemical Theory and Computation* **2019**, *15*, 1827–1840.

(147) Singraber, A.; Morawietz, T.; Behler, J.; Dellago, C. *Journal of Chemical Theory and Computation* **2019**, *15*, 3075–3092.

(148) Sun, G.; Sautet, P. *Journal of Chemical Theory and Computation* **2019**, *15*, Publisher: American Chemical Society, 5614–5627.

(149) Zhang, Y.-R.; Zhao, Z.-Z.; Xue, C.; Gao, F.; Wang, Y.-N. *Journal of Physics D: Applied Physics* **2019**, *52*, 295204.

(150) Basher, A. H.; Hamada, I.; Hamaguchi, S. *Japanese Journal of Applied Physics* **2020**, *59*, Publisher: IOP Publishing, 090905.

(151) Basher, A. H.; Krstić, M.; Fink, K.; Ito, T.; Karahashi, K.; Wenzel, W.; Hamaguchi, S. *Journal of Vacuum Science & Technology A* **2020**, *38*, Publisher: American Vacuum Society, 052602.

(152) Basher, A. H.; Krstić, M.; Takeuchi, T.; Isobe, M.; Ito, T.; Kiuchi, M.; Karahashi, K.; Wenzel, W.; Hamaguchi, S. *Journal of Vacuum Science & Technology A* **2020**, *38*, Publisher: American Vacuum Society, 022610.

(153) Chi, H.; Curnan, M. T.; Li, M.; Andolina, C. M.; Saidi, W. A.; Veser, G.; Yang, J. C. *Physical Chemistry Chemical Physics* **2020**, *22*, 2738–2742.

(154) Chowdhury, A. J.; Yang, W.; Abdelfatah, K. E.; Zare, M.; Heyden, A.; Terejanu, G. A. *Journal of Chemical Theory and Computation* **2020**, *16*, Publisher: American Chemical Society, 1105–1114.

(155) George, S. M. *Accounts of Chemical Research* **2020**, *53*, Publisher: American Chemical Society, 1151–1160.

(156) Gerrits, N.; Juaristi, J. I.; Meyer, J. *Physical Review B* **2020**, *102*, 155130.

(157) Mullins, R.; Kondati Natarajan, S.; Elliott, S. D.; Nolan, M. *Chemistry of Materials* **2020**, *32*, Publisher: American Chemical Society, 3414–3426.

(158) Rahm, J. M.; Erhart, P. *J. Open Source Softw.* **2020**, *5*, 1944.

(159) Sang, X.; Chang, J. P. *Journal of Vacuum Science & Technology A* **2020**, *38*, Publisher: American Vacuum Society, 042604.

(160) Sang, X.; Chen, E.; Chang, J. P. *Journal of Vacuum Science & Technology A* **2020**, *38*, Publisher: American Vacuum Society, 042603.

(161) Sang, X.; Xia, Y.; Sautet, P.; Chang, J. P. *Journal of Vacuum Science & Technology A* **2020**, *38*, Publisher: American Vacuum Society, 043005.

(162) Sang, X.; Xia, Y.; Sautet, P.; Chang, J. P. *Journal of Vacuum Science & Technology A* **2020**, *38*, Publisher: American Vacuum Society, 043005.

(163) Schran, C.; Behler, J.; Marx, D. *Journal of Chemical Theory and Computation* **2020**, *16*, 88–99.

(164) Staub, R.; Steinmann, S. N. *The Journal of Chemical Physics* **2020**, *152*, 024124.

(165) Zhu, W.; Gong, H.; Han, Y.; Zhang, M.; van Duin, A. C. T. *The Journal of Physical Chemistry C* **2020**, *124*, Publisher: American Chemical Society, 12512–12520.

(166) Fischer, A.; Routzahn, A.; George, S. M.; Lill, T. *Journal of Vacuum Science & Technology A* **2021**, *39*, Publisher: American Vacuum Society, 030801.

(167) Konh, M.; Janotti, A.; Teplyakov, A. *The Journal of Physical Chemistry C* **2021**, *125*, Publisher: American Chemical Society, 7142–7154.

(168) Kunze, S.; Tănase, L. C.; Prieto, M. J.; Grosse, P.; Scholten, F.; de Souza Caldas, L.; van Vörden, D.; Schmidt, T.; Cuenya, B. R. *Chemical Science* **2021**, *12*, 14241–14253.

(169) Li, L.; Chen, G.; Zheng, H.; Meng, W.; Jia, S.; Zhao, L.; Zhao, P.; Zhang, Y.; Huang, S.; Huang, T.; Wang, J. *Nature Communications* **2021**, *12*, 3863.

(170) Li, M.; Curnan, M. T.; Gresh-Sill, M. A.; House, S. D.; Saidi, W. A.; Yang, J. C. *Nature Communications* **2021**, *12*, 2781.

(171) Musil, F.; Grisafi, A.; Bartók, A. P.; Ortner, C.; Csányi, G.; Ceriotti, M. *Chemical Reviews* **2021**, *121*, Publisher: American Chemical Society, 9759–9815.

(172) Sharma, V.; Elliott, S. D.; Blomberg, T.; Haukka, S.; Givens, M. E.; Tuominen, M.; Ritala, M. *Chemistry of Materials* **2021**, *33*, Publisher: American Chemical Society, 2883–2893.

(173) Sheil, R.; Martirez, J. M. P.; Sang, X.; Carter, E. A.; Chang, J. P. *The Journal of Physical Chemistry C* **2021**, *125*, 1819–1832.

(174) Xia, Y.; Sautet, P. *Chemistry of Materials* **2021**, *33*, 6774–6786.

(175) Thompson, A. P.; Aktulga, H. M.; Berger, R.; Bolintineanu, D. S.; Brown, W. M.; Crozier, P. S.; in 't Veld, P. J.; Kohlmeyer, A.; Moore, S. G.; Nguyen, T. D.; Shan, R.; Stevens, M. J.; Tranchida, J.; Trott, C.; Plimpton, S. J. *Computer Physics Communications* **2022**, *271*, 108171.

(176) Kim, Y.; Chae, S.; Ha, H.; Lee, H.; Lee, S.; Chae, H. *Applied Surface Science* **2023**, *619*, 156751.

(177) Wang, P.; Castelli, M.; Fang, F. *Materials Science in Semiconductor Processing* **2023**, *153*, 107169.

(178) US 6,245,670 B1 - Method for filling a dual damascene opening having high aspect ratio to minimize electromigration failure | RPX Insight.

(179) Cheung, R.; Lopatin, S. US 6,245,670 B1 - Method for filling a dual damascene opening having high aspect ratio to minimize electromigration failure | RPX Insight, U.S. Patent 6245670B1, 2001.