Plasma Oxidation of Copper: Molecular Dynamics Study with Neural Network Potentials

Yantao Xia[±‡] and Philippe Sautet[±‡,]

‡*Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, CA 90095, USA*

Department of Chemistry and Biochemistry, University of California, Los Angeles, CA 90095, USA

E-mail: xyttyxy@ucla.edu; sautet@ucla.edu

Abstract

The formation of thin oxide films is of significant scientific and practical interest. In particular, the semiconductor industry is interested in developing a plasma atomic layer etching process to pattern copper, replacing the dual Damascene process. Using a nonthermal oxygen plasma to convert the metallic copper into copper oxide, followed by a formic acid organometallic reaction to etch the copper oxide, this process has shown great promise. However, the current process is not optimal because the plasma oxidation step is not self-limiting, hampering the degree of thickness control. In the present study, a neural network potential for the binary interaction between copper and oxygen is developed and validated against first principles calculations. This potential covers the entire range of potential energy surfaces of metallic copper, copper oxides, atomic oxygen, and molecular oxygen. The usable kinetic energy ranges from 0 eV to 20 eV. Using this potential, the plasma oxidation of copper surfaces was studied with large-scale molecular dynamics at atomic resolution, with an accuracy approaching

that of the first principle calculations. An amorphous layer of CuO is formed on Cu, with thicknesses reaching 2.5 nm. Plasma is found to create an intense local heating effect that rapidly dissipates across the thickness of the film. The range of this heating effect depends on the kinetic energy of the ions. A higher ion energy leads to a longer range, which sustains faster-than-thermal rates for longer periods of time for the oxide growth. Beyond the range of this agitation, the growth is expected to be limited to the thermally activated rate. High-frequency, repeated ion impacts result in a microannealing effect that leads to a quasicrystalline oxide beneath the amorphized layer. The crystalline layer slows down oxide growth. Growth rate is fitted to the temperature gradient due to ion-induced thermal agitations, to obtain an apparent activation energy of 1.0 eV. A strategy of lowering the substrate temperature and increasing plasma power is proposed as being favorable for more self-limited oxidation.

Keywords

atomic layer etching; plasma simulation; machine learning; neural network potential; molecular dynamics; metal oxidation

Introduction

After replacing aluminum decades ago, copper became the universal material of choice for interconnects in integrated circuits.[1] 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.[2] 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.[3]

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).[4] To etch metallic copper, a specific variant, the plasma-thermal ALE process can be used. It is composed of two steps (Figure 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.[5] 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.[6]

Copper oxide is one of the enabler materials of modern technology. Stable under ambient conditions with an optimal band gap of $\sqrt{}$ 2.0 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,[7], copper oxidation is now considered somewhat of a model system to understand formation of thin metal oxides in general.[8]

The oxidation of low-index copper surfaces has been extensively studied under surface science conditions.[9–13] 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,[14] 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.*[15]

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.[16] The particles used in the corresponding experimental process (the "target process") have kinetic energies in the order of 10 eV,[17,18] 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.

To the authors' 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,[19–23] and are performed with classical interatomic potentials (*e.g.* Tersorff-Brenner[24,25] & Stillinger-Weber[26]). 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[27], MEAM[28], MEAM-Qeq[29]) and the bond order method (COMB[30,31], ReaxFF[32,33]). 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 ph-

ysisorbed molecules, and to some extent the oxide), and the ionic bonds (in the oxide) interactions at the same time. Recently ReaxFF[34] and COMB[31,35] 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)[36] and the Gaussian approximation potentials (GAP)[37] 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.

In the present paper, a high-dimensional neural network potential (HDNNP) for the binary interaction of copper and oxygen is trained on the density functional theory (DFT) potential energy surface. It will be shown that 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. The interaction potential is applied to plasma oxidation of copper to predict the influence of processing conditions such as substrate temperature, plasma power, and chamber pressure. Finally, suggestions on improving the existing experimental process are given on the basis of the influences found.

Results and Discussion

With the trained and validated potential, large-scale molecular dynamics simulations are performed. Figure 2 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 2b). 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

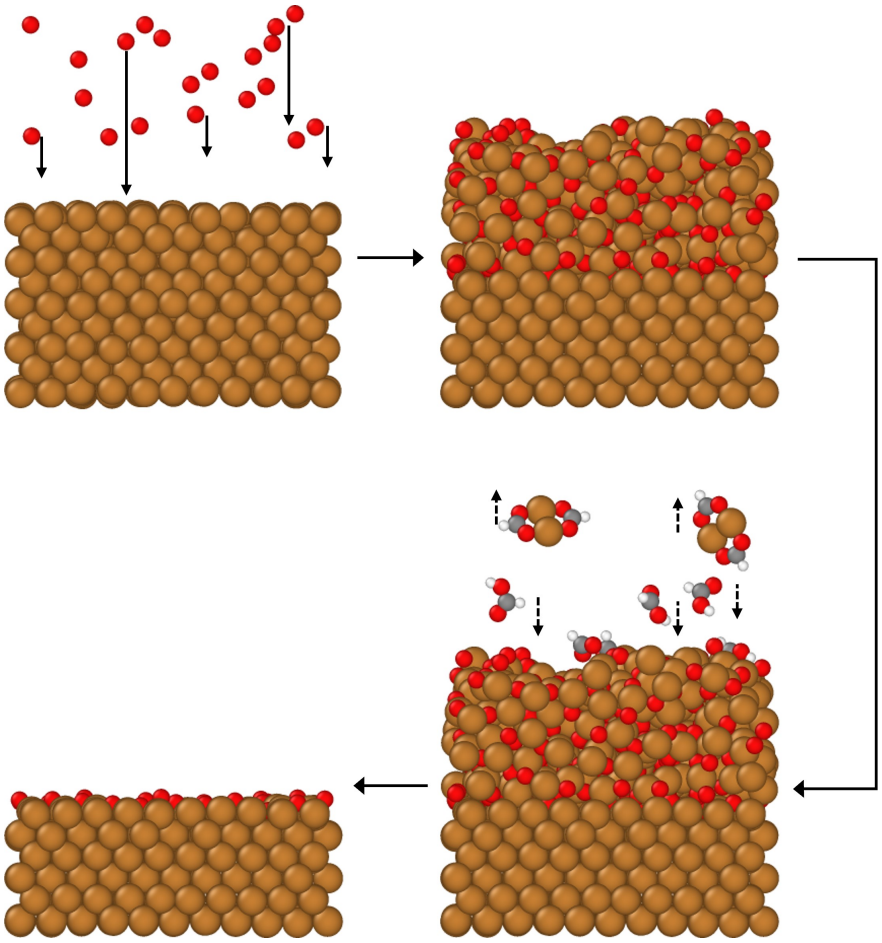

Figure 1: Overall process of plasma-thermal ALE on Cu metal. **a)**: Starting with a pristine surface, reactive oxygen neutrals and ions are generated by the plasma. The ions are accelerated by the plasma sheath toward the surface, indicated by the long arrows. The neutrals, much larger in number, diffuse 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.

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.[5] Apart from the velocity, the charge is not explicitly modeled since most of the ions are neutralized by Auger electrons before reaching the surface.[16] 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

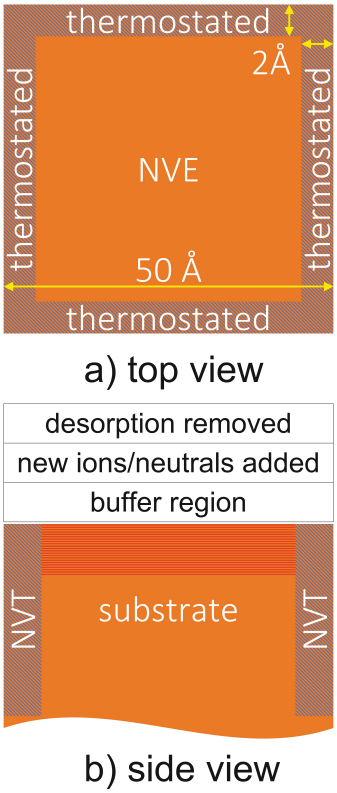

Figure 2: Overall setup of MD simulations. The production simulations used $(20 * 20 * 50)$ slabs. The schematic here is not drawn to scale. **(a)**: the lateral cross section is divided into thermostated edges and non-thermostated centers; **(b)**: the vertical space in the simulation box is divided into the desorption removal, new atoms addition, buffer, and the substrate regions

trend is clear: Higher plasma power and higher pressure correspond to more ions of higher energy, at the possible expense of decreased directionality.

In Figure 3, 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 3, $t = 56\,\mathrm{ps}$). 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 3, $t = 170\,\mathrm{ps}$, 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 3, the oxide growth between 1587 ps and 4892 ps only amounts to approximately two atomic

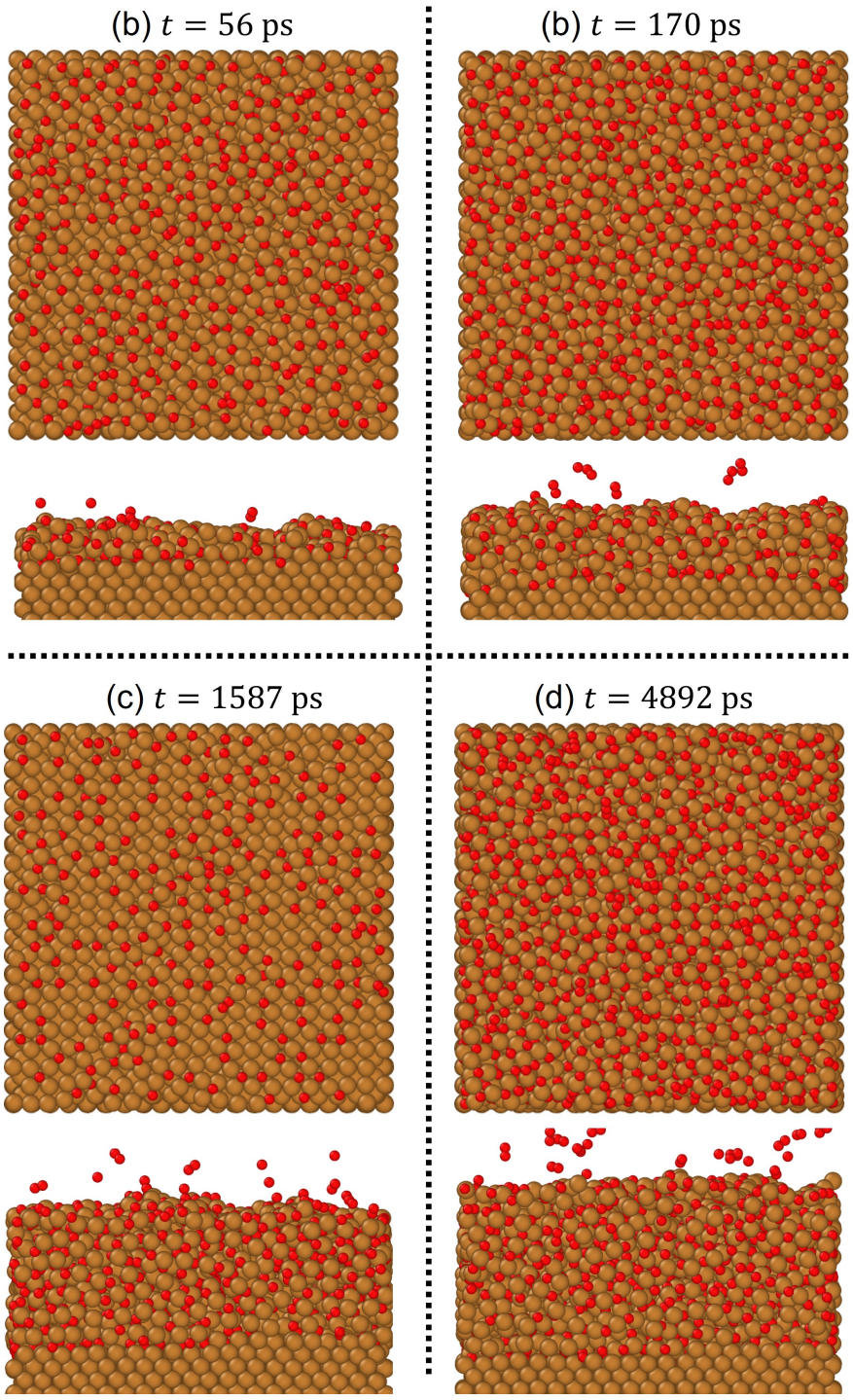

Figure 3: Snapshots along the oxidation trajectory of Cu(100) slab at four different points in time. **(a), (b)**: the very early stage where the growth is determined by the availability of oxygen (controlled by the neutral to ion ratio since ions are deposited at fixed step intervals). Between **(a)** 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.

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 4a) shows the growth of oxide thickness with time, and Figure 4b) shows the amount of oxygen incorporated into the slab. The details of determining this thickness from the MD trajectory are given in Section S2.1. The recombined O_2 molecules can be seen in the side views shown in Figure 3. 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 3, from the first snapshot, it can be

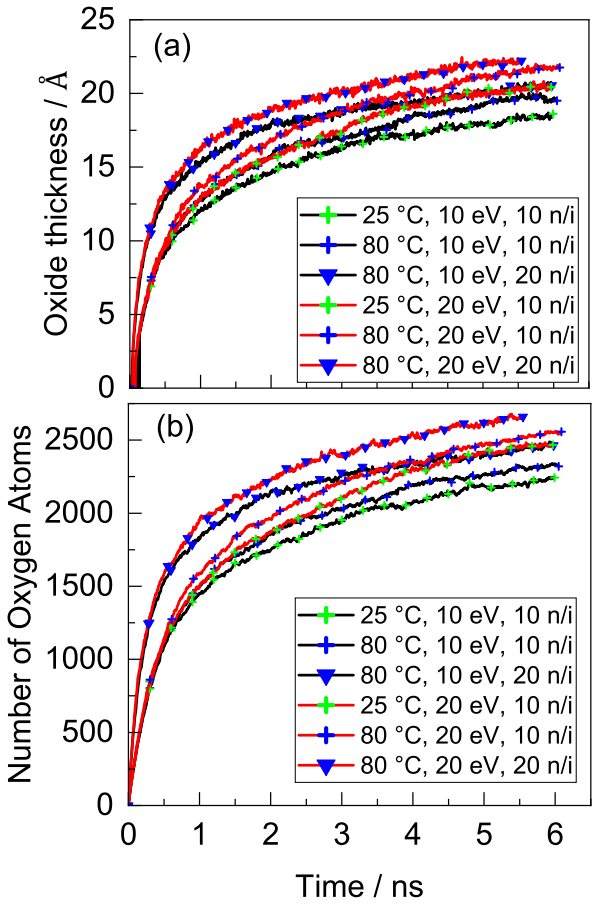

Figure 4: Thickness and oxygen incorporation as functions of simulation time for all six conditions explored. The growth curves start as two distinct strands, where the grouping is determined by the neutral to ion ratio (indicated by the shape of the marker, cross vs triangle). The strands quickly splits 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.

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.

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 3 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 $\sqrt{3}\,\mathrm{ps}$. In Figure 5a-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 $f_{\text{cut}} = 1 * 10^{-5}$ 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 Section S2.4). 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 5a) to 20 eV (Figure 5b) 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 5c, 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 $\sqrt{}$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.

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 will 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

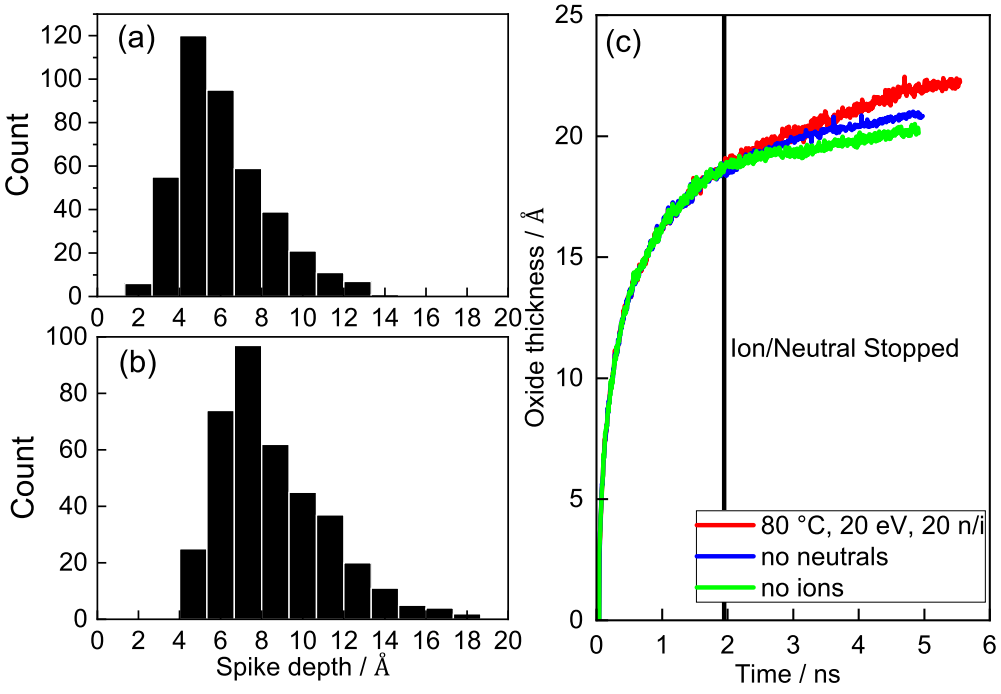

Figure 5: Distribution of the projected range of thermal agitation on copper oxide slabs. The oxide slab is obtained by first taking an thick oxide slab resulting from the long-time simulation. Oxygen ions are added to the slab with different random seeds determining the lateral position. 500 replicas 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

the deposition rate is much higher than what is likely under experimental pressures (see the comparison in Section S2.2).

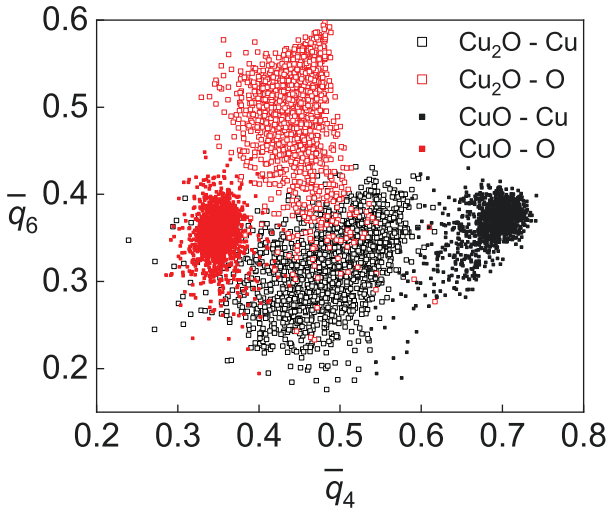

Figure 6: Bond orientational order parameters for bulk CuO and Cu_2O, used as references for distinguishing local environments of the oxide slab obtained from MD. The distinguishing feature is similar $\bar{q}_4$ parameters and different $\bar{q}_6$ for CuO, but the opposite for Cu_2O.

Voronoi tessellation averaged bond orientational order (BOO) parameters is used to study the crystal structure of the copper oxide layers.[38,39] 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 Section S2.3. Shown in Figure 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 $\sqrt{}0.40$, but different $\bar{q}_4$ values at $\sqrt{}0.35$ and $\sqrt{}0.70$, corresponding to O and Cu atoms, respectively. On the other hand, Cu_2O has two clusters with similar $\bar{q}_4$ values at $\sqrt{}0.50$, but different $\bar{q}_6$ values centered at $\sqrt{}0.5$ and $\sqrt{}0.3$, respectively, for O and Cu. The spread is due to thermal fluctuations as a result of the high-temperature equilibrium MD.

In Figure 7, the calculated BOO parameter are plotted against their vertical co-ordinate (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 S6. The graphs used in Figure 7 are

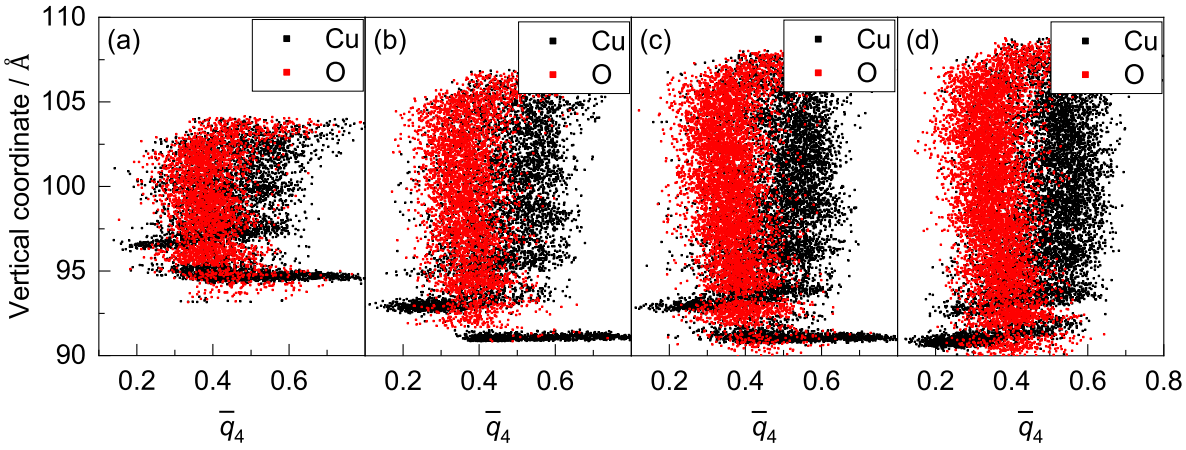

Figure 7: Bond orientational order parameter $\bar{q}_4$ for a range of oxide thicknesses. The structures used are taken from the $T_s = 353\,\mathrm{K}, E_k = 20\,\mathrm{eV}, n/i = 10$. The z-coordinates of atoms are plotted against the $\bar{q}_4$ values, showing a gradual horizontal separation of 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 S6 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.

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 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 suggests that the resulting oxide after long exposure to oxygen plasma will have the structure and composition of CuO. This result is consistent with experimental X-ray photoelectron spectroscopy study,[5] although in that case the Cu_2O is not ruled out. The $\bar{q}_6$ result suggests no Cu_2O structure is present (Figure S6). The overall stoichiometry of the oxide film is close to 1 (see Figure 10 for local atomic concentrations), which agrees with the analysis of the local environment.

In Figure 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 8a, BOO analysis similar to that of Figure 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 8a show clear clustering in the z-coordinate. The structure itself shows quasicrystalline CuO regions within $\sqrt{5}$ 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 S2, confirming that the local environment closely resembles CuO. The z-coordinates of the clusters are labelled in Figure 8. The distances between successive Cu and O clusters are calculated to be $(1.27 \sim 0.02)$ 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 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 4. 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 7 and 8 does not indicate a distinct structure, but is a manifestation of the unreliable BOO parameters on surfaces due to the tendency of the Voronoi tessellation definition of the nearest neighbors to overestimate the number of neighbors[40]. 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 (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 (2)). Therefore, the 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 5 were reused, since the deposition frequencies matched those in long-time MD (Figure 3). The oxide layer is divided into 20 slices

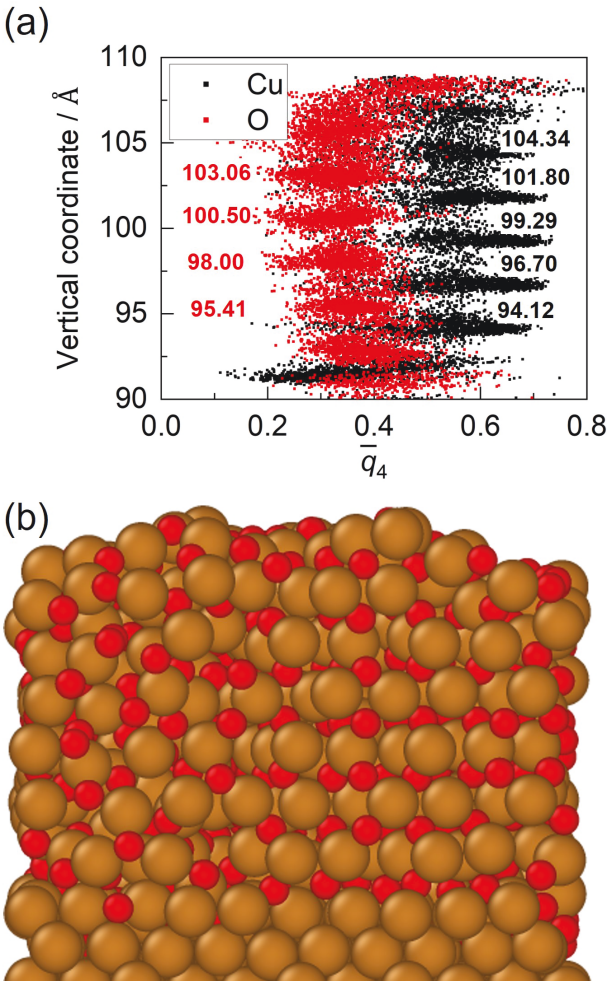

Figure 8: Bond orientational order parameter (BOO) on $(10 * 10)$ CuO (100) slabs and the corresponding structure. At 1/4 of the surface area, the ion impact frequency were maintained at 10000 steps per ion. The neutral to ion ratio were kept at 10. This led to a 4-fold increase in effective ion and 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.

along the z-coordinate of the atoms, and the temperature distribution in each slice was calculated separately within each slice for each atom. Figure S7 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 (3)). Optimization is performed on the objective function in Equation (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,

$$\frac{dL}{dt} = A \times \exp\left(\frac{E_A}{k_B T}\right) \tag{1}$$

$$\bar{r}(L) = \frac{1}{N}\sum_{i}^{N} \frac{dL}{dt} \tag{2}$$

$$L^{\mathrm{rp}}(t_i) = L^{\mathrm{rp}}(t_{i-1}) + \bar{r}(L^{\mathrm{rp}}(t_{i-1}))\delta t \tag{3}$$

$$F = \sqrt{\sum_{t=t_0,\ldots,t_i,\ldots,t_f} (L^{\mathrm{rp}}(t_i) - L^{\mathrm{md}}(t_i))^2} \tag{4}$$

Additional details and intermediate results are given in Section S2.5.

Figure 9 shows the result of this optimization. Fitting E_A and A to the growth curve of $T_s = 353\,\mathrm{K}$, $E_K = 10\,\mathrm{eV}$, 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 is not captured by the simple model. Long-time behavior is emphasized in Equation (4) due to the larger number of points in this region, leading to the reproduced curve (Figure 9, 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 9, red curves), suggesting that the rate-limiting process is independent of the kinetic energy of the ion.

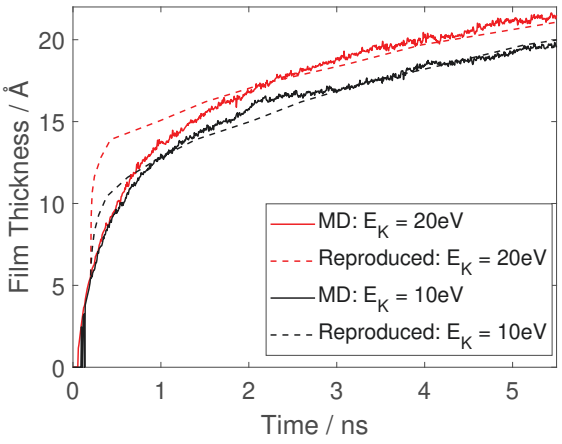

Figure 9: Comparison of the fitted thickness curve to the MD result. The apparent activation energy E_A and pre-exponential factor A is fitted to the thickness-time relation of the $T_s = 353\,\mathrm{K}, E_k = 10\,\mathrm{eV}, n/i = 10$ trajectory from the temperature distribution obtained using $E_k = 10\,\mathrm{eV}$ ions. The resulting optimal values of $E_A = 1.01\,\mathrm{eV}$ and $A = 5.63 * 10^7\,\mathrm{nm/ns}$ 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.

The fitted activation barrier can be compared to a direct calculation of the diffusion 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[41,42], and computationally calculated to be around 0.5 eV[43]. 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[44], 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. In Figure S10 we provide NNP-based nudged elastic band calculation of the diffusion barriers, showing excellent agreement to literature DFT values, again demonstrating the accuracy of the model.

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 10 as a function of thickness. It suggests that the film has a porous region within $\surd$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 $\surd$8 Å above the oxidation front (as defined in Section S2.1), does not have local environments 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 4. It also explains the trend observed when the neutrals are removed from the simulation (Figure 5c), where the growth seems to follow the case with both ions and neutrals for $\surd$2 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 5c.

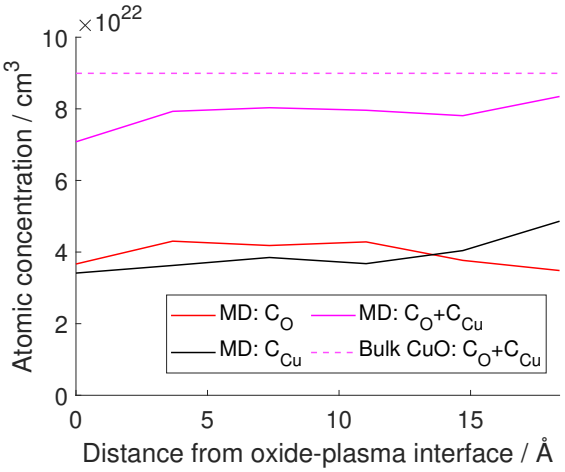

Figure 10: Atomic concentration in the oxide film. The film is taken from the $T_s = 353\,\mathrm{K}, E_k = 20\,\mathrm{eV}, n/i = 10$ trajectory at time $t = 5.54\,\mathrm{ns}$. The result here indicate a porous region where the total atomic density is lower, a dense region where the Cu and O concentration 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.,[45] 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.[5] 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 will need to be bridged.

Important differences exist between the work of Kunze *et. al.* and the target ex-

perimental 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.[15] 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, will further diminish CuO. 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 will 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.[46]

To ensure the results is not particular to the Cu (100) slab, similar long-time MD simulations were performed on the Cu (111) surface. Figure S12 shows the evolution

of the chosen rectangular unit cell on the Cu(111) surface, and Figure S11 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 S11). 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.

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 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. 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, $\surd$ 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.

Methods

All electronic structure calculations are performed with the density functional theory as implemented in the Vienna ab initio Simulation Package (VASP).[47–49] The electron-ion interactions are treated using the projector augmented wave (PAW) method[50] and the valence one-electron functions are developed on a basis set of plane waves. The PBE exchange-correlation functional[51] 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 S1.1) 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.[52] 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.

The molecular dynamics simulations were performed using the Large-scale Atomic / Molecular Massively Parallel Simulator (LAMMPS).[53], in conjunction with the ReaxFF package[54] and the NNP interface to the n2p2 neural network potential li-

brary.[55]. Details about MD simulation (*e.g.* choice of time step size, thermostatting, time integration, *etc.*) are given in Section S1.4. The standalone ReaxFF package was used to reparameterize the inner-wall parameters to stabilize the very initial ReaxFF simulations.

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`).[56]. The procedure is discussed in detail in Section S1.5. Additional computational details on the architecture of the network are given in Section S1.2.

There are several types of structures used in the training set generation and production MD simulations. Throughout this paper, the convension $(w * l * t)$ 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) $(3 * 3 * 5)$ surface which is subsequently oxidized. A small subset of training MD simulation started with $(3 * 3 * 10)$ Cu(100) surfaces. For these surfaces, a $(3 * 3 * 1)$ k-space mesh was used. Furthermore, bulk oxide structures Cu_2O and CuO were used with $(2 * 2 * 2)$ supercells of conventional unit cell for both. The reciprocal space was sampled with $(3 * 3 * 3)$ and $(5 * 5 * 3)$ meshes for Cu_2O and CuO, respectively. For the production MD (simulations used to obtain the results), the Cu(100) surface was expanded to a $(20 * 20)$ surface cell in all cases except for results used in Figure 8, which considers a $(10 * 10)$ surface cell. Additionally, a $(8^{\dagger}\ \overline{3} * 8^{\dagger}\ \overline{3})$ supercell for the Cu(111) surface was used in Section S4. In these simulations, the number of layers is not of interest because the oxide never exhausted the copper layers.

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 parame-

terization reported in the literature[57] 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. Details of this modification can be found in Section S1.5.1. 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, $\sqrt{1000}$ 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 ($< 5\,\mathrm{meV/atom}$) 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 S1.5.2). Data cleaning is done before and after DFT calculations, based on geometric and energetic criteria, respectively (see Section S1.5.3).

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 $\surd$ 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 details are given in Section S1.5.4.

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 11 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 ($d > \surd 3.5$ Å), the quintet is $\surd$ 2 eV more stable than the singlet, in accordance with the well-known singlet-triplet energy difference of O_2.[58] The adjacent triplet molecules adopt an antiferromagnetic (AFM) coupling to form an overall singlet ground state complex, qualitatively consistent with the literature.[59] 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.[60–62] 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 will be 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.[63,64]

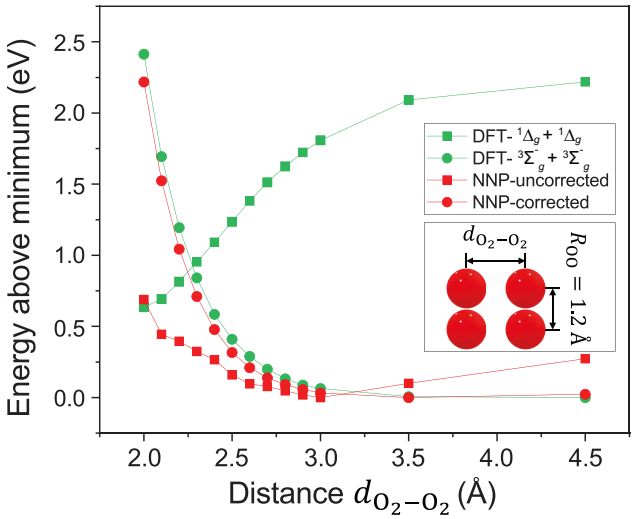

Figure 11: Comparison of the potential energy surfaces of O_2 complex in the "H" configuration, shown in inset. The corrected NNP is shown to follow the triplet state PES.

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 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 S4).

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 discussion of these sub-datasets is relegated to the Supplementary Information. The overall energy error 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 S5. It can be seen that all of the datasets have similarly low energy errors. The average adsorption energies on representative configurations from 3 different phases of oxidation are shown in Figure 12. Two types of slab structures are used: the $(6* 6* 3)$ slab and the $(3* 3* 10)$ 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

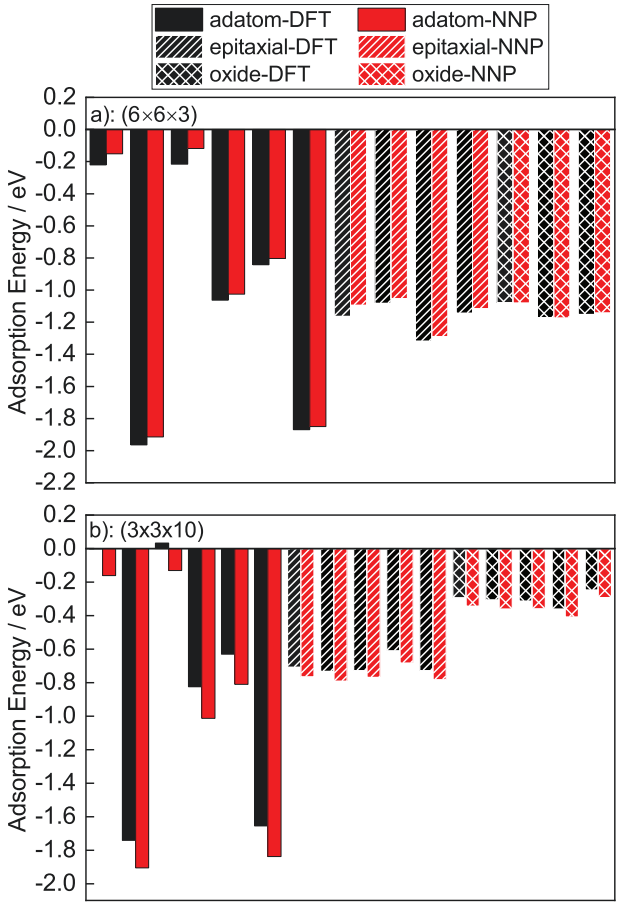

Figure 12: Comparison of average adsorption energies on realistic test cases including single adatom adsorption, epitaxial oxide on metal, and fully oxidized slabs. **(a)**: result on a $(6 * 6 * 3)$, thin Cu(100) slab. **(b)**: result on a $(3 * 3 * 10)$ Cu(100) slab.

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 S1.3) 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 13a), vertical (Figure 13b) and flat (Figure 13c) O_2 molecules. The structures are illustrated in Figure S5. 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.

Supporting Information

Additional information including attempts at using higher level of theory, quantification of errors introduced by the adiabatic PES assumption, details of the MD setup, reparametrization of ReaxFF as the starting point for dataset generation, methods of quantifying oxide film thickness, comparison of theoretical and realistic deposition rates, characterizations of the oxide film composition, intermediate results for fitting the diffusion apparent activation energy, and details of direct calculation of the diffusion barrier.

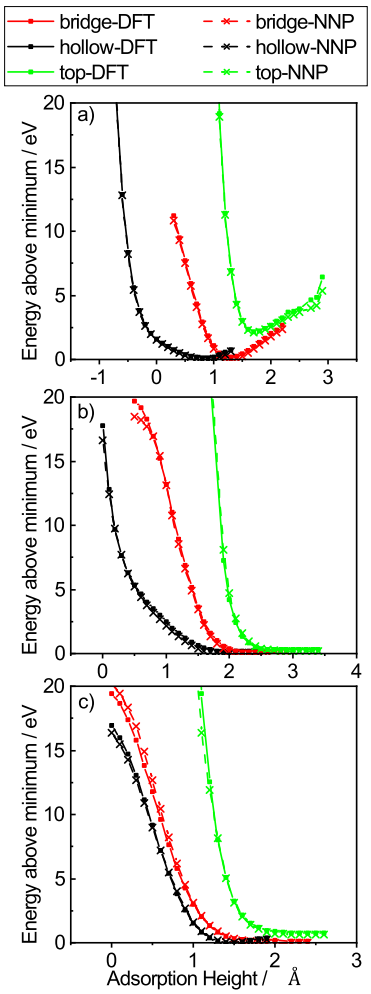

Figure 13: Comparison of potential energy surfaces on short-distance, high-energy structures likely present in impact events. **(a)**: a single O atom; **(b)**: a vertical O_2 molecule; **(c)**: horizontal O_2 molecule. The energies are plotted against the distance above the adsorption sites on a Cu (100) surface.

Acknowledgements

This work was funded by NSF, CBET division, Grant No. 2212981. The authors gratefully acknowledge the support from UCLA Institute of Digital Research and Education (IDRE) for computational resources on the UCLA Hoffman2 cluster. This work used the Extreme Science and Engineering Discovery Environment Bridges-2 cluster at Pittsburgh Computing Center. These XSEDE resources are accessed through allocation TG-CHE170060. XSEDE is supported by National Science Foundation Grant Number ACI-1548562.

Associated Content

This manuscript was previously made public on GitHub[65].

References

1. Edelstein, D.; Heidenreich, J.; Goldblatt, R.; Cote, W.; Uzoh, C.; Lustig, N.; Roper, P.; McDevitt, T.; Motsiff, W.; Simon, A.; Dukovic, J.; Wachnik, R.; Rathore, H.; Schulz, R.; Su, L.; Luce, S.; Slattery, J. Full copper wiring in a sub-0.25 um CMOS ULSI technology. International Electron Devices Meeting. IEDM Technical Digest. 1997; pp 773–776.

2. Rosenberg, R.; Edelstein, D. C.; Hu, C.-K.; Rodbell, K. P. Copper Metallization for High Performance Silicon Technology. *Annual Review of Materials Science* **2000**, *30*, 229–262.

3. 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; `https://insight.rpxcorp.com/patent/US6245670B1`.

4. Carver, C. T.; Plombon, J. J.; Romero, P. E.; Suri, S.; Tronic, T. A.; Turkot, R. B. Atomic Layer Etching: An Industry Perspective. *ECS Journal of Solid State Science and Technology* **2015**, *4*, N5005–N5009.

5. Sheil, R.; Martirez, J. M. P.; Sang, X.; Carter, E. A.; Chang, J. P. Precise Control of Nanoscale Cu Etching via Gas-Phase Oxidation and Chemical Complexation. *The Journal of Physical Chemistry C* **2021**, *125*, 1819–1832.

6. Altieri, N. D.; Chen, J. K.-C.; Minardi, L.; Chang, J. P. Review Article: Plasma–surface interactions at the atomic scale for patterning metals. *Journal of Vacuum Science & Technology A* **2017**, *35*, 05C203.

7. Hellman, A. Mechanism for Limiting Thickness of Thin Oxide Films on Aluminum. *Chemistry of Materials* **2016**,

8. Gattinoni, C.; Michaelides, A. Atomistic details of oxide surfaces and surface oxidation: the example of copper and its oxides. *Surface Science Reports* **2015**, *70*, 424–447.

9. Zhou, G.; Luo, L.; Li, L.; Ciston, J.; Stach, E. A.; Yang, J. C. Step-Edge-Induced Oxide Growth During the Oxidation of Cu Surfaces. *Physical Review Letters* **2012**, *109*, 235502.

10. Zhu, Q.; Zou, L.; Zhou, G.; Saidi, W. A.; Yang, J. C. Early and transient stages of Cu oxidation: Atomistic insights from theoretical simulations and in situ experiments. *Surface Science* **2016**, *652*, 98–113.

11. Zhu, Q.; Saidi, W. A.; Yang, J. C. Step-Induced Oxygen Upward Diffusion on Stepped Cu(100) Surface. *The Journal of Physical Chemistry C* **2015**, *119*, 251–261.

12. Li, M.; Curnan, M. T.; Gresh-Sill, M. A.; House, S. D.; Saidi, W. A.; Yang, J. C. Unusual layer-by-layer growth of epitaxial oxide islands during Cu oxidation. *Nature Communications* **2021**, *12*, 2781.

13. Chi, H.; Curnan, M. T.; Li, M.; Andolina, C. M.; Saidi, W. A.; Veser, G.; Yang, J. C. In situ environmental TEM observation of two-stage shrinking of Cu2O islands on Cu(100) during methanol reduction. *Physical Chemistry Chemical Physics* **2020**, *22*, 2738–2742.

14. Yang, J. C.; Kolasa, B.; Gibson, J. M.; Yeadon, M. Self-limiting oxidation of copper. *Applied Physics Letters* **1998**, *73*, 2841–2843.

15. Choudhary, S.; Sarma, J. V. N.; Pande, S.; Ababou-Girard, S.; Turban, P.; Lepine, B.; Gangopadhyay, S. Oxidation mechanism of thin Cu films: A gateway towards the formation of single oxide phase. *AIP Advances* **2018**, *8*, 055114.

16. Neyts, E. C.; Brault, P. Molecular Dynamics Simulations for Plasma-Surface Interactions. *Plasma Processes and Polymers* **2017**, *14*, 1600145.

17. Zhang, Y.-R.; Zhao, Z.-Z.; Xue, C.; Gao, F.; Wang, Y.-N. Ion energy and angular distributions in planar Ar/O2 inductively coupled plasmas: hybrid simulation and experimental validation. *Journal of Physics D: Applied Physics* **2019**, *52*, 295204.

18. Wang, Y.; Olthoff, J. K. Ion energy distributions in inductively coupled radio-frequency discharges in argon, nitrogen, oxygen, chlorine, and their mixtures. *Journal of Applied Physics* **1999**, *85*, 6358–6365.

19. Athavale, S. D.; Economou, D. J. Molecular dynamics simulation of atomic layer etching of silicon. *Journal of Vacuum Science & Technology A* **1995**, *13*, 966–971.

20. Gou, F.; Neyts, E.; Eckert, M.; Tinck, S.; Bogaerts, A. Molecular dynamics simulations of Cl+ etching on a Si(100) surface. *Journal of Applied Physics* **2010**, *107*, 113305.

21. Rauf, S.; Sparks, T.; Ventzek, P. L. G.; Smirnov, V. V.; Stengach, A. V.; Gaynullin, K. G.; Pavlovsky, V. A. A molecular dynamics investigation of fluorocarbon based layer-by-layer etching of silicon and SiO2. *Journal of Applied Physics* **2007**, *101*, 033308.

22. 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. Realistic Atomistic Structure of Amorphous Silicon from Machine-Learning-Driven Molecular Dynamics. *The Journal of Physical Chemistry Letters* **2018**, *9*, 2879–2885, Publisher: American Chemical Society.

23. Graves, D. B.; Brault, P. Molecular dynamics for low temperature plasma–surface interaction studies. *Journal of Physics D: Applied Physics* **2009**, *42*, 194011.

24. Tersoff, J. New empirical approach for the structure and energy of covalent systems. *Physical Review B* **1988**, *37*, 6991–7000.

25. Brenner, D. W. Empirical potential for hydrocarbons for use in simulating the chemical vapor deposition of diamond films. *Physical Review B* **1990**, *42*, 9458–9471.

26. Stillinger, F. H.; Weber, T. A. Computer simulation of local order in condensed phases of silicon. *Physical Review B* **1985**, *31*, 5262–5271.

27. Daw, M. S.; Baskes, M. I. Embedded-atom method: Derivation and application to impurities, surfaces, and other defects in metals. *Physical Review B* **1984**, *29*, 6443–6453.

28. Baskes, M. I. Modified embedded-atom potentials for cubic materials and impurities. *Physical Review B* **1992**, *46*, 2727–2742.

29. Rappe, A. K.; Goddard, W. A. Charge equilibration for molecular dynamics simulations. *The Journal of Physical Chemistry* **1991**, *95*, 3358–3363.

30. Shan, T.-R.; Devine, B. D.; Kemper, T. W.; Sinnott, S. B.; Phillpot, S. R. Charge-optimized many-body potential for the hafnium/hafnium oxide system. *Physical Review B* **2010**, *81*, 125328.

31. Liang, T.; Shan, T.-R.; Cheng, Y.-T.; Devine, B. D.; Noordhoek, M.; Li, Y.; Lu, Z.; Phillpot, S. R.; Sinnott, S. B. Classical atomistic simulations of surfaces and heterogeneous interfaces with the charge-optimized many body (COMB) potentials. *Materials Science and Engineering: R: Reports* **2013**, *74*, 255–279.

32. van Duin, A. C. T.; Dasgupta, S.; Lorant, F.; Goddard, W. A. ReaxFF: A Reactive Force Field for Hydrocarbons. *The Journal of Physical Chemistry A* **2001**, *105*, 9396–9409.

33. 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. The ReaxFF reactive force-field: development, applications and future directions. *npj Computational Materials* **2016**, *2*, 15011.

34. Jeon, B.; Sankaranarayanan, S. K.; van Duin, A. C.; Ramanathan, S. Influence of surface orientation and defects on early-stage oxidation and ultrathin oxide growth on pure copper. *Philosophical Magazine* **2011**, *91*, 4073–4088.

35. Devine, B.; Shan, T.-R.; Cheng, Y.-T.; McGaughey, A. J. H.; Lee, M.; Phillpot, S. R.; Sinnott, S. B. Atomistic simulations of copper oxidation and Cu/Cu2O interfaces using charge-optimized many-body potentials. *Physical Review B* **2011**, *84*, 125308, Publisher: American Physical Society.

36. Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Physical Review Letters* **2007**, *98*, 146401.

37. Bartók, A. P.; Payne, M. C.; Kondor, R.; Csányi, G. Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons. *Physical Review Letters* **2010**, *104*, 136403.

38. Steinhardt, P. J.; Nelson, D. R.; Ronchetti, M. Bond-orientational order in liquids and glasses. *Physical Review B* **1983**, *28*, 784–805, Publisher: American Physical Society.

39. Mickel, W.; Kapfer, S. C.; Schröder-Turk, G. E.; Mecke, K. Shortcomings of the bond orientational order parameters for the analysis of disordered particulate matter. *The Journal of Chemical Physics* **2013**, *138*, 044501.

40. Staub, R.; Steinmann, S. N. Parameter-free coordination numbers for solutions and interfaces. *The Journal of Chemical Physics* **2020**, *152*, 024124.

41. Magnuson, H.; Frisk, K. *Self-diffusion and impurity diffusion of hydrogen, oxygen, sulphur and phosphorus in copper*; 2014; pp 14–16, Swedish Nuclear Fuel and Waste Management Company.

42. Narula, M. L.; Tare, V. B.; Worrell, W. L. Diffusivity and solubility of oxygen in solid copper using potentiostatic and potentiometric techniques. *Metallurgical Transactions B* **1983**, *14*, 673–677.

43. Li, C.; Zhang, P.; Wang, J.; Boscoboinik, J. A.; Zhou, G. Tuning the Deoxygenation of Bulk-Dissolved Oxygen in Copper. *The Journal of Physical Chemistry C* **2018**, *122*, 8254–8261.

44. Li, L.; Chen, G.; Zheng, H.; Meng, W.; Jia, S.; Zhao, L.; Zhao, P.; Zhang, Y.; Huang, S.; Huang, T.; Wang, J. Room-temperature oxygen vacancy migration induced reversible phase transformation during the anelastic deformation in CuO. *Nature Communications* **2021**, *12*, 3863.

45. Kunze, S.; Tănase, L. C.; Prieto, M. J.; Grosse, P.; Scholten, F.; de Souza Caldas, L.; van Vörden, D.; Schmidt, T.; Cuenya, B. R. Plasma-assisted oxidation of Cu(100) and Cu(111). *Chemical Science* **2021**, *12*, 14241–14253.

46. Xia, Y.; Sautet, P. Thermodynamics of Atomic Layer Etching Chemistry on Copper and Nickel Surfaces from First Principles. *Chemistry of Materials* **2021**, *33*, 6774–6786.

47. Blöchl, P. E. Projector augmented-wave method. *Phys. Rev. B* **1994**, *50*, 17953–17979.

48. Kresse, G.; Furthmüller, J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B* **1996**, *54*, 11169–11186.

49. Kresse, G.; Furthmüller, J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Computational Materials Science* **1996**, *6*, 15 – 50.

50. Kresse, G.; Joubert, D. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B* **1999**, *59*, 1758–1775.

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

52. Haynes, W. M. , *CRC Handbook of Chemistry and Physics, 97th Edition*, 97th ed.; CRC Press: Boca Raton, FL, 2016; pp 5–11, 5–29, 5–30.

53. 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. LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales. *Computer Physics Communications* **2022**, *271*, 108171.

54. Aktulga, H. M.; Fogarty, J. C.; Pandit, S. A.; Grama, A. Y. Parallel reactive molecular dynamics: Numerical methods and algorithmic techniques. *Parallel Computing* **2012**, *38*, 245–259.

55. Singraber, A.; Behler, J.; Dellago, C. Library-Based LAMMPS Implementation of High-Dimensional Neural Network Potentials. *Journal of Chemical Theory and Computation* **2019**, *15*, 1827–1840.

56. Singraber, A.; Morawietz, T.; Behler, J.; Dellago, C. Parallel Multistream Training of High-Dimensional Neural Network Potentials. *Journal of Chemical Theory and Computation* **2019**, *15*, 3075–3092.

57. Zhu, W.; Gong, H.; Han, Y.; Zhang, M.; van Duin, A. C. T. Development of a Reactive Force Field for Simulations on the Catalytic Conversion of C/H/O Molecules on Cu-Metal and Cu-Oxide Surfaces and Application to Cu/CuO-Based Chemical Looping. *The Journal of Physical Chemistry C* **2020**, *124*, 12512–12520.

58. Schweitzer, C.; Schmidt, R. Physical Mechanisms of Generation and Deactivation of Singlet Oxygen. *Chemical Reviews* **2003**, *103*, 1685–1758.

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

60. Zagidullin, M. V.; Pershin, A. A.; Azyazov, V. N.; Mebel, A. M. Luminescence of the (O2(a1Δg))2 collisional complex in the temperature range of 90-315 K: Experiment and theory. *The Journal of Chemical Physics* **2015**, *143*, 244315.

61. Bussery-Honvault, B.; Veyret, V. Comparative studies of the lowest singlet states of (O2)2 including ab initio calculations of the four excited states dissociating into O2(1Δg)+O2(1Δg). *The Journal of Chemical Physics* **1998**, *108*, 3243–3248.

62. Hernández-Lamoneda, R.; Hernández, M.; Campos-Martínez, J. A new singlet ab initio potential energy surface for studying vibrational relaxation in O2(v)+O2 collisions. *Chemical Physics Letters* **2003**, *368*, 709–716.

63. Obata, M.; Nakamura, M.; Hamada, I.; Oda, T. Implementation of van der Waals Density Functional Approach to the Spin-Polarized System: Interaction Potential between Oxygen Molecules. *Journal of the Physical Society of Japan* **2013**, *82*, 093701.

64. Aquilanti, V.; Ascenzi, D.; Bartolomei, M.; Cappelletti, D.; Cavalli, S.; de Castro Vítores, M.; Pirani, F. Quantum Interference Scattering of Aligned Molecules:

Bonding in O4 and Role of Spin Coupling. *Physical Review Letters* **1999**, *82*, 69–72.

65. Xia, Y.; Sautet, P. GitHub, https://github.com/xyttyxy/cuonnppaper/, Accessed on 2022-11-18.