



Multiscale Modeling of Defect Phenomena in Platinum Using Machine Learning of Force Fields

JAMES CHAPMAN¹ and RAMPI RAMPRASAD^{1,2}

1.—School of Materials Science and Engineering, Georgia Institute of Technology, North Ave NW, Atlanta, GA 30332, USA. 2.—e-mail: rampi.ramprasad@mse.gatech.edu

Computational methodologies have been critical to our understanding of defects at nanometer scales. These methodologies have been dominated by two classes: quantum mechanics (QM)-based methods and semiempirical/classical methods. The former, while accurate and versatile, are time consuming, while the latter are efficient but limited in versatility and transferability. Recently, machine learning (ML) methods have shown initial promise in bridging these two limitations due to their accuracy and flexibility. In this work, the true capability of ML methods is explored by simulating defects in platinum over several length/time scales. We compare our results with density functional theory (DFT) for atomic-level defect behavior and with experiments for nanolevel behavior. We also compare our predictions with several classical potentials. This work aims to showcase the length/time scales attainable using ML, as well as the complexity they are capable of capturing, demonstrating that these methodologies may be effectively used, in the future, to bridge experiments and QM methods.

INTRODUCTION

Atomistic computational techniques have been instrumental in the exploration of material defects at both atomic and nanometer length scales.^{1–7} These methods have historically fallen into two broad categories: QM-based methods, e.g., density functional theory (DFT),^{8,9} and semiempirical methods, e.g., the embedded atom method.^{10–16} While both groups have been used to study a multitude of materials phenomena,^{17–19} they both suffer from serious shortcomings. QM methods, while able to capture properties at a high level of fidelity, are computationally expensive, which severely restricts both the time and length scales that can be reliably accessed through such simulations. Semiempirical/classical methods significantly reduce this cost burden and allow for the exploration of length and time scales that are not attainable with QM. However, as such methods are parameterized to tackle specific problems, they do not attain the same

level of versatility as QM methods, showing a steep decline in accuracy away from their respective reference data.

To this end, data-driven machine learning (ML) methods have shown promise as a reliable alternative, bridging the gap in cost, accuracy, and transferability.^{20–28} Unlike their counterparts, ML methods rely on functional forms, and parameterizations of these functional forms, that are derived from statistics, rather than physics. The accuracy of these models will still decline during extrapolative predictions, and in this regard, they are no better or worse than semiempirical/classical methodologies. However, ML approaches offer a number of advantages over these methods; for instance, their predictive capabilities may be iteratively improved in a systematic manner.^{1,29–37}

Recently, ML methods have been deployed to study a variety of defect properties in elemental metals,^{38–40} including our previous work on platinum.^{37,41} While the ML methods presented in those works have shown great promise in their ability to accurately predict atomic-level properties (at the same time and length scales as their respective reference data), with respect to DFT, phenomena

requiring much larger time ($> \text{ns}$) and length scales ($> \text{nm}$) have not been widely explored, or validated (e.g., against experiments).

In this manuscript, we build upon our previous ML force field for Pt^{37,41} by iteratively improving its training data to predict a multitude of defect properties in platinum over several length and time scales. Atomic-level properties, such as grain boundary and surface energies, as well as adatom and dimer binding and adsorption energies on several surfaces are calculated to show the capabilities of the models for accurately capturing the complexity and diversity of atomic-level defect properties. Grain-coarsening simulations are also performed, and compared with experimental observations.⁴²

The remainder of this manuscript is organized as follows: We first begin by providing a brief overview of the adaptive generalizable neighborhood informed (AGNI) methodology. Second, we discuss our strategy to iteratively improve of these models over previous protocols. We then discuss several atomic-level properties of platinum for surfaces, line defects, planar defects, and combinations of defects. Finally, we discuss the temperature dependence of the grain size evolution via molecular dynamics simulations. The compilation of work presented in this manuscript aims to reveal the potential of ML methodologies for the study of materials phenomena over large time and length scales, bridging the gap between QM, semiempirical/classical, and experimental methodologies.

COMPUTATIONAL DETAILS

AGNI Workflow

The AGNI platform consists of several key steps, regardless of the property (atomic forces, potential energy, stresses, and electronic structure) being predicted: (1) generation of a diverse set of reference data, (2) numerical encoding of local/structural geometric information (fingerprinting), (3) training a ML model given some subset of the reference data, and (4) employing the final ML models in an MD engine, capable of simulating the dynamic time evolution of atomistic processes. In the following sections we provide a brief explanation of steps 1, 2, and 3, and we refer the reader to our previous works for a more thorough understanding.^{29–31,36,37,43,44}

Reference Data Generation

A comprehensive set of reference data, summarized in Table I, was prepared for Pt in an accurate and uniform manner in order to minimize the numerical noise intrinsic to atomistic calculations. All reference data were obtained using the Vienna *ab initio* simulation package (VASP).^{45–49} The Perdew–Burke–Ernzerhof (PBE) functional⁵⁰ was used to calculate the electronic exchange–correlation interaction. Projector augmented wave (PAW) potentials⁵¹ and plane-wave basis functions up to

a kinetic energy cutoff of 500 eV were used. All projection operators (involved in the calculation of the nonlocal part of the PAW pseudopotentials) were evaluated in reciprocal space to ensure further precision. Monkhorst–Pack⁵² \mathbf{k} -point meshes were carefully calibrated for each atomic configuration to ensure numerical convergence in both energy and atomic forces.

Fingerprinting Atomic Configurations

A hierarchical representation of an atom's local coordination was created to capture geometric information that is mapped directly to properties such as the total potential energy, atomic forces, and stresses. This hierarchy aims to capture unique components of an atom's local neighborhood with features resembling scalar, vector, and tensor quantities. The functional forms of all atomic-level fingerprint components are defined as follows:^{43,44}

$$S_{i,k} = c_k \sum_{j \neq i} \exp \left[-\frac{1}{2} \left(\frac{r_{ij}}{\sigma_k} \right)^2 \right] f_{\text{cut}}(r_{ij}), \quad (1)$$

$$V_{i,x;k} = c_k \sum_{j \neq i} \frac{r_{ij}^\alpha}{r_{ij}} \exp \left[-\frac{1}{2} \left(\frac{r_{ij}}{\sigma_k} \right)^2 \right] f_{\text{cut}}(r_{ij}), \quad (2)$$

$$T_{i,\{x,\beta\};k} = c_k \sum_{j \neq i} \frac{r_{ij}^\alpha r_{ij}^\beta}{r_{ij}^2} \exp \left[-\frac{1}{2} \left(\frac{r_{ij}}{\sigma_k} \right)^2 \right] f_{\text{cut}}(r_{ij}), \quad (3)$$

with r_i and r_j being the Cartesian coordinates of atoms i and j , and $r_{ij} = |\mathbf{r}_j - \mathbf{r}_i|$. α and β represent any of the three directions x , y , or z . The σ_k values control the width of the Gaussian functions, and are determined via a grid-based optimization process.³¹ The damping function $f_{\text{cut}}(r_{ij}) = \frac{1}{2} [\cos(\frac{\pi r_{ij}}{R_{\text{cut}}}) + 1]$, which smoothly decays towards zero, has a cutoff radius R_{cut} chosen to be 8 Å. c_k is a normalization constant given by $(\frac{1}{\sigma_k \sqrt{2\pi}})^3$ (for the force model, this normalization constant was set to 1).

To learn rotationally invariant properties, such as the total potential energy, a separate step is required to map the atomic fingerprints described above to rotationally invariant versions, defined as^{37,43,44}

$$V_{i,k} = \sqrt{(V_{i,x;k})^2 + (V_{i,y;k})^2 + (V_{i,z;k})^2}, \quad (4)$$

$$T'_{i,k} = T_{i,\{x,x\};k} T_{i,\{y,y\};k} + T_{i,\{x,x\};k} T_{i,\{z,z\};k} + T_{i,\{y,y\};k} T_{i,\{z,z\};k} - (T_{i,\{x,y\};k})^2 - (T_{i,\{x,z\};k})^2 - (T_{i,\{y,z\};k})^2, \quad (5)$$

and

$$T''_{i,k} = \det(T_{i,\{\alpha,\beta\};k}). \quad (6)$$

Table I. Summary of reference dataset prepared for platinum ML model generation

Defect Type	Systems	Temperature (K)	ML Trained on
Defect-free	Bulk (w/o strain)	300, 1000, 2000	Yes
Strained bulk	Bulk (w/o strain \pm 7%)	300, 1000, 2000	Yes
Monovacancy	Bulk with one vacancy	0 (NEB), 1000, 1500, 2000	Yes
Divacancy	Bulk with divacancy	0 (NEB), 1000, 1500, 2000	Yes
Clustered vacancies	Bulk with four vacancies	1000, 1500, 2000	Yes
Grain boundary	Assorted GBs between \sum 3 and \sum 25	0 (relaxed), 300, 1000, 1500	Yes
Grain boundary + vacancy	Assorted GBs between \sum 3 and \sum 25	0 (relaxed), 0 (NEB)	No
Surface	Slabs corresponding to GB planes	0, 300, 600, 1000, 1500	Yes
Surface + adatom/dimer	Adatoms and dimers on (111), (110), and (100)	0 (relaxed), 0 (NEB)	No

The data are divided into subsets based on the type of defect present. $T = 0$ K designates either nudged elastic band (NEB) calculations or geometry relaxations, while $T > 0$ K indicates MD calculations. The final column (“ML Trained on”) indicates whether or not the final ML models included that row of data in their respective training sets

For global properties, such as potential energy and stress tensor, the atomic fingerprints described above are not sufficient, as they cannot map directly to a global property, only per-atom quantities. Therefore, a second process mapping the atomic fingerprints to a single, structural fingerprint is required.⁴³ In this work, the ML models that learn the potential energy, as well as the stress tensor, employ such a procedure. This mapping involves summing the atomic fingerprints, over all atoms in the system, and taking the first moment of this sum, $M^1(X)$, with X representing the sum of the atomic fingerprints. The first moment can be interpreted as the average atomic environment of the system. The final forms of all the fingerprints for energy and stresses are presented in Supplementary Table S1.

Machine Learning

After the final fingerprint forms have been crafted, and a subset of our total reference data has been sampled, we employ Kernel ridge regression (KRR) to establish a mapping between our fingerprints and the atomic forces, potential energy, and total stress tensor. This learning scheme employs a similarity-based nonlinear kernel to create a mapping between the reference fingerprints and the desired property using a functional form described as^{1,29–32,37}

$$P_X = \sum_Y \alpha_Y \exp \left[-\frac{1}{2} \left(\frac{d_{XY}}{\sigma} \right)^2 \right]. \quad (7)$$

Here the summation runs over the number of reference environments Y in a given model’s training set. As each model is trained separately, there is no restriction that each model must have the same number of training points. P symbolizes the desired

property (total potential energy, stress tensor components, or atomic forces), with X being the fingerprint of a new configuration. d_{XY} represents the Euclidean distance between fingerprints X and Y , calculated within the feature hyperspace, and σ is a length-scale parameter. During the model’s training phase, the regression weights α_Y and the length scale σ are determined via a regularized objective function, which is optimized through a fivefold cross-validation process. At the end of the model construction/optimization process, three independent ML models for energy, forces, and stresses exist (see Supplementary Table S3 for all ML hyperparameters).

Simulation Details

Three embedded atom method (EAM) potentials, henceforth referred to as EAM-B⁴ EAM-Z,⁵ and EAM-F⁶ (due to their respective authors), were chosen for this work and used for comparison with the ML platform presented in this work. All EAM potentials were chosen for their ability to accurately capture a variety of bulk properties for Pt with respect to experimental evidence. Details regarding the level of theory for all DFT calculations used in this work can be found in the reference data section, and all details regarding the ML scheme used can be found in the previous three sections. This ML scheme has been benchmarked against both EAM and DFT, for the calculation of energy, forces, and stresses, and is approximately five orders of magnitude faster than DFT, but roughly two orders of magnitude slower than EAM. Throughout this work, three distinctive classes of calculations are used to study materials properties over several length scales: (1) geometry optimizations, (2) nudged elastic band calculations, and (3) molecular dynamics simulations.

Geometry optimizations were performed using both VASP (for DFT) and the large-scale atomic/molecular massively parallel simulator (LAMMPS) package⁵³ (for EAM and ML). These calculations were used to gather information such as the grain boundary energy, surface energy, work of separation energy, vacancy formation energy for in-plane vacancies of grain boundaries, and the relaxed structure of both edge and screw dislocations. All grain boundary configurations used in this work were constructed using Aimgb software.⁵⁴ In this work, grain boundary energy was calculated as

$$\gamma_{\text{GB}} = \frac{E_{\text{GB}} - n_{\text{GB}}E_0}{2A_{\text{GB}}}, \quad (8)$$

with n_{GB} being the number of atoms in the grain boundary structure, A_{GB} being the area of the grain boundary plane, and E_0 being the cohesive energy. The surface energy was calculated using the same formula, substituting n_{GB} for n_{Surf} , and A_{GB} for A_{Surf} . A factor of 2 is used in both cases to account for the two identical surfaces, or grain boundaries, that are contained in each unit cell. The work of separation energy was calculated as

$$W_{\text{sep}} = 2\gamma_{\text{Surf}} - \gamma_{\text{GB}}. \quad (9)$$

The grain boundary vacancy formation energy, with respect to both a bulk configuration as well as a pristine grain boundary environment, was also calculated. In both cases, the vacancy was placed at the center of the boundary plane. For boundaries that span multiple layers, the vacancy was placed as close to the midpoint (normal to the boundary) as possible. The vacancy formation energy with respect to bulk was calculated as

$$E_{\text{vac}} = E_{\text{GB+vac}} - n_{\text{GB+vac}}E_0, \quad (10)$$

with $n_{\text{GB+vac}}$ being the number of atoms in the grain boundary structure containing a single vacancy, and $E_{\text{GB+vac}}$ being the corresponding energy of the structure. When using the pristine grain boundary as the reference, Eq. 10 is used, but E_0 is substituted for E_{GB} , where E_{GB} is the per-atom energy of the grain boundary system without a vacancy.

Adsorption energies for both a single adatom as well as a dimer were calculated on the Pt (111), (110), and (100) surfaces, using an 80-, 144-, and 180-atom slab, respectively. The binding energy for a dimer on each of these surfaces was also calculated. The adsorption energy was calculated as

$$E_{\text{ads}} = -\frac{1}{N}(E_{\text{slab+adsorbate}} - (E_{\text{slab}} + NE_{\text{atom}})). \quad (11)$$

N is defined as the number of adsorbates in the system. $E_{\text{slab+adsorbate}}$ is defined as the energy of the slab with the adsorbate bonded to the surface, E_{slab} is the energy of just the slab (no adsorbate), and E_{atom} is the energy of a single adsorbate atom in a box. The binding energy of a dimer is also defined as

$$E_{\text{bind}} = E_{\text{separate}} - E_{\text{together}}, \quad (12)$$

where E_{separate} is defined as the energy of a system containing a slab with the two atoms in the dimer placed as far away from each other as possible. E_{together} is defined as the energy of the slab with the dimer bonded to the surface, and is equivalent to $E_{\text{slab+adsorbate}}$ for the case of the dimer.

In all scenarios, both the ionic positions and the cell volume were allowed to change. Electronic convergence terminated at an energy difference of 10^{-4} eV, and ionic relaxations were considered converged at an energy difference of 10^{-2} eV, for all calculations.

Nudged elastic band (NEB) calculations, along with the climbing image formalism,⁴⁹ were employed to determine the minimum-energy pathway of a single vacancy diffusing both along, and away from, the boundary plane, of the $\Sigma 3111$ grain boundary, as well as the activation energies of several adatom diffusion mechanisms on the (111), (110), and (100) surfaces. Similarly to the geometry optimizations, both VASP (DFT) and LAMMPS (EAM and ML) were used. In all scenarios, both the ionic positions and the cell volume were allowed to change. Electronic convergence terminated at an energy difference of 10^{-4} eV, and ionic relaxations were considered converged at an energy difference of 10^{-2} eV, for all calculations.

Molecular dynamics simulations, using LAMMPS, were performed to study how grain sizes are affected by temperature. A $51 \times 51 \times 51$ supercell containing 508,971 atoms was used. The initial distribution of grains was created using the Voronoi tessellation.⁵⁵ *NPT* simulations, run for 1 ns at $P = 0$, were used to equilibrate the supercell volume at a given temperature. Simulations were performed between $T = 300$ K and $T = 700$ K, to align with experimental results.⁴² Previous computational work indicates that 1 ns is sufficient to allow for equilibration among grain sizes at the temperatures and system size considered in this work.^{56,57}

RESULTS AND DISCUSSION

Iterative Improvement of ML Models

In our previous work, three independent ML models (for energy, atomic forces, and the total stress tensor) were created to study elastic, diffusive, and thermal properties of platinum. However, these systems contained either strained bulk, or point defect configurations. More complex defects, such as planar and line defects, and combinations of defect classes were not present in the ML model's respective training sets. Before performing simulations of these new systems, model predictions of potential energy, atomic forces, and the total stress tensor were compared with the reference DFT data. Parity plots, shown in Fig. 1 (top), indicate that, while the previous ML models perform well for

grain boundary configurations, they cannot make reliable predictions on surface environments. Therefore, these systems must be added to the three model's training sets before performing any simulations.

Using an iterative improvement scheme developed for our ML force models,³⁶ failed configurations were continuously added to all three ML models' training sets until statistical metrics reached convergence. From Fig. 1 (bottom), a clear improvement in all ML models can be seen, and statistics for the final retrained ML models are presented in Supplementary Table S2. With converged statistical metrics showing significant improvement in the model performance, when encountering previously disparate domains in the configuration space, the ML models can now be reliably deployed to these regions during real simulations.

Edge and Screw Dislocations

Edge and screw dislocations play a crucial role in the plasticity and fracture of metals.^{58–61} Therefore, accurately predicting the geometry of dislocation

lines is of paramount importance. To this end, we studied $\frac{1}{2}[110]$ edge and screw dislocations. As Pt is a face-centered cubic (fcc) metal, the most favorable slip system is $\frac{1}{2}\langle 110 \rangle 111$,⁶² We employ the dislocation extraction algorithm (DXA) to determine the geometric information around the dislocation core. Using a $1 \times 1 \times 3$ supercell (due to the requirements of the DXA algorithm,⁶³ and the reduction in the z -direction required for the system sizes needed to perform a DFT relaxation) of the relaxed dislocation system, the dislocation type, dislocation line length, and the dislocation's Burgers vector were determined.

When considering the edge dislocation, upon relaxation, the dislocation type determined by the DXA analysis was $\frac{1}{2}\langle 110 \rangle$ for all levels of theory used in this work. The dislocation line length was calculated as 14.61 Å for DFT, and 14.22 Å, 13.89 Å, and 13.91 Å for ML, EAM-B, EAM-F, and EAM-Z, respectively. The Burgers vector for all models was calculated as $\frac{1}{2}[0\bar{1}1]$. Upon relaxation, the initial dislocation core split into two cores that

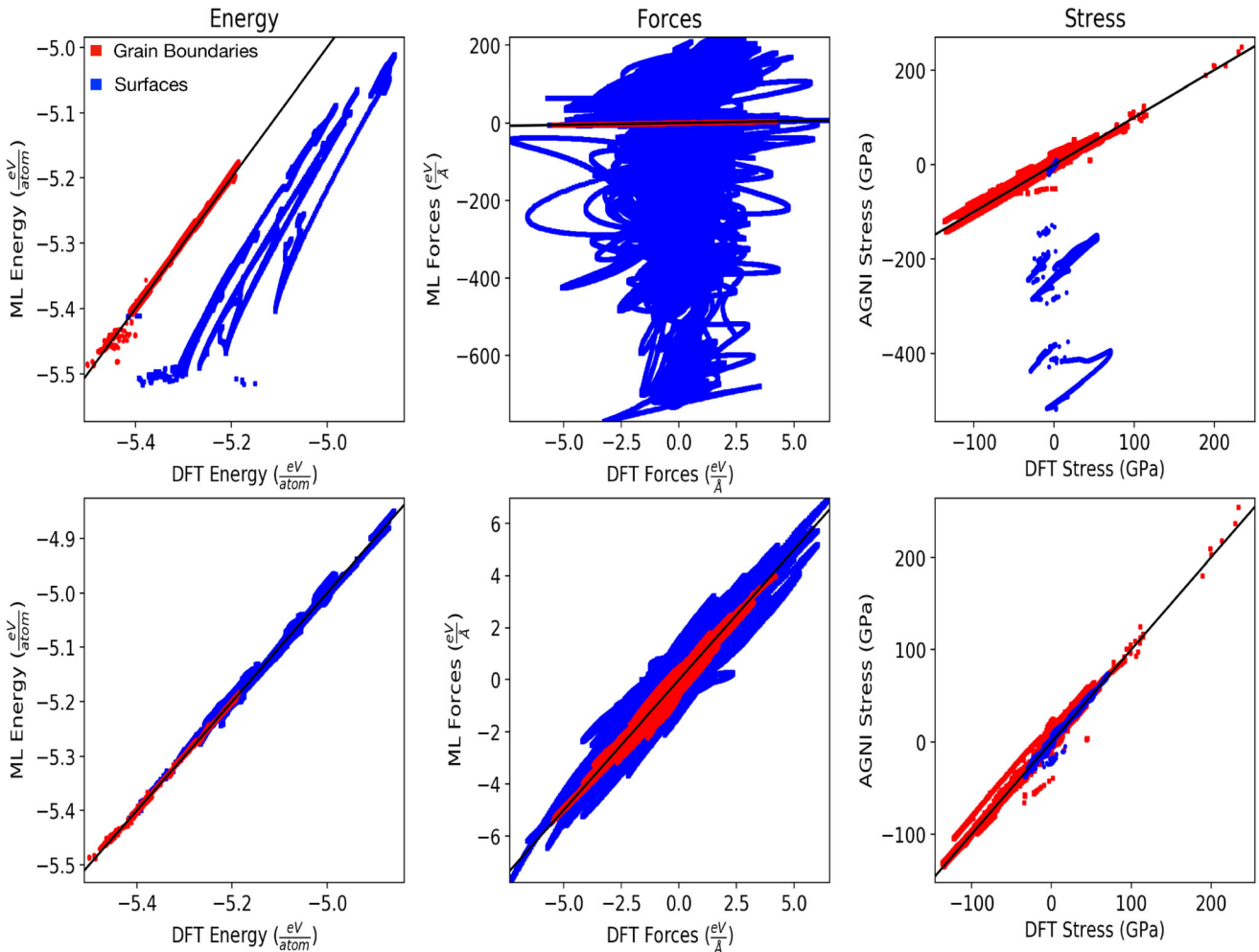


Fig. 1. Parity plots of the energy, forces, and stresses predicted on the total set of grain boundary (red) and surface (blue) reference data before (top) and after (bottom) the ML models iterative improvement procedure.

migrated away from each other along the dislocation's Burgers vector. The final core centers were located 9.31 Å apart for DFT and 9.11 Å, 15.03 Å, 18.05 Å, and 18.05 Å for ML, EAM-B, EAM-F, and EAM-Z, respectively. The agreement between ML and DFT can be seen in both the dislocation line length as well as the separation distance between dislocation cores. While the calculated dislocation line lengths for all EAM potentials will not necessarily agree with DFT due to differences in the equilibrium lattice parameters, there does exist substantial disagreement in the core separation distance between the EAM models used in this work and DFT.

When considering the screw dislocation, upon relaxation, the dislocation type determined by the DXA analysis was $\frac{1}{2} \langle 110 \rangle$ for all models used. The dislocation line length was calculated as 8.439 Å for DFT, and 8.58 Å, 8.37 Å, 8.27 Å, and 8.28 Å for ML, EAM-B, EAM-F, and EAM-Z respectively. The Burgers vector for all models was calculated as $\frac{1}{2} [0\bar{1}1]$. Unlike the edge dislocation, the initial screw dislocation core remained intact after relaxation for all levels of theory. While there is no discernible difference between ML/EAM and DFT for the relaxed screw dislocation structure, there exists a substantial discrepancy between the relaxed EAM and DFT edge dislocation structure. This disparity is alleviated when using the ML models, which show near-perfect agreement with DFT.

Clean Surfaces

In this work, we calculate properties of clean surfaces, such as the surface energy and the interlayer relaxation difference for a variety of low-index surfaces. Figure 3b shows all the surface energy calculations for ML and the three EAM potentials with respect to the corresponding DFT prediction. Statistically, our ML models have a root-mean-square error (RMSE) of 0.08 J/m² with respect to the DFT grain boundary energies. EAM-B, EAM-Z, and EAM-F yield RMSE values of 0.10 J/m², 0.62 J/m², and 0.05 J/m², respectively. Here, EAM-F performs slightly better than our ML models, though both models perform equally well as the surface complexity increases (see Supplementary Table S4).

The interlayer relaxation difference, δd_{12} and δd_{23} , was calculated for the (111), (110), and (100) surfaces. Table II provides the calculated values, along with experimental predictions where available.^{64,65} For the (111) surface, δd_{12} is calculated as +0.92% for DFT, in good agreement with experiments, which indicate a positive relaxation difference. While our ML models indicate a smaller percent change than both DFT and experiments, they do predict a positive interlayer relaxation difference. This is in contrast with all EAM potentials, which predict a negative interlayer relaxation difference. Experimental predictions indicate that

there is no substantial difference between the second and third layers, though both DFT and ML yield a small negative value for δd_{23} . However all EAM potentials predict a positive difference.

For the (110) surface, δd_{12} is calculated as -15.92% by DFT, in good agreement with experiments. Our ML models again predict a value roughly half that of the DFT value, but correctly capture its negative nature. Among the EAM potentials, there exists a substantial spread, with EAM-F performing extremely well, EAM-Z performing comparably to our ML models, and EAM-B showing significant deviation from both DFT and experiments. For δd_{23} , DFT deviates significantly from experiments, with DFT predicting a positive interlayer relaxation difference and experiments indicating a negative difference. As our ML models are trained on DFT data, they are expected to follow its trend. Our ML models are in near-perfect agreement with DFT, and all levels-of-theory indicate that a positive δd_{23} occurs.

For the (100) surface, there is a significant spread among experiments regarding δd_{12} , with some predicting a positive difference and others a negative difference. DFT, ML, and EAM-B fall within the possible values indicated by experiments, while both EAM-F and EAM-Z fall outside of the experimental spread. As no experimental values exist for δd_{23} , we can only compare our ML models with DFT, indicating good agreement. Both EAM-B and EAM-F deviate significantly from DFT, though EAM-Z is in good agreement, though we would like to emphasize that the true value is unknown.

Surfaces with Adsorbates

In this section, we discuss the introduction of adatoms onto the (111), (110), and (100) surfaces. First we consider a single adatom on each surface. Upon relaxing each system, the adatom adsorption energy, and distance between adatom and surface atoms is calculated, and can be found in Table II. Our ML models, as well as all EAM potentials, agree well with DFT with regards to the bond distance, when considering the differences in equilibrium lattice parameter. However, our ML models, while predicting the exact trend in adsorption energy with respect to DFT, yields values approximately 1 eV lower than DFT. For all surfaces, there exists a spread among EAM values, but the trend in adsorption energy with respect to DFT is captured exactly.

The activation energy for adatom diffusion is also considered, calculated via the NEB method, and can be found in Table II (for plot see Supplementary Fig. S1). For all surfaces, two mechanisms are considered: (1) hop and (2) two-atom exchange. For the (111) surface hop and exchange, our ML models are in excellent agreement with both DFT and experiments.^{66,67} All EAM potentials significantly underestimate the transition state energy of the hop

Table II. Properties of low-index platinum surfaces

Property	Exp.	DFT	ML	EAM-B	EAM-F	EAM-Z
(111)						
δd_{12}	$+1.1 \pm 0.5$	+0.92	+0.24	-1.29	-4.11	-2.32
δd_{23}	0.0	-0.71	-0.54	+1.41	+1.27	+0.47
d_{adatom}		2.57	2.54	2.56	2.35	2.35
$E_{1,\text{hop}}$	0.26 ± 0.02	0.32	0.28	0.09	0.03	0.05
$E_{1,\text{exc}}$		1.70	1.75	1.70	0.82	1.90
$E_{1,\text{ads}}$		4.45	3.75	4.39	4.44	3.97
$E_{2,\text{ads}}$		4.61	3.89	4.62	4.90	4.37
$E_{2,\text{bind}}$		0.54	0.60	0.42	0.94	0.82
(110)						
δd_{12}	-18.0 ± 5.0	-15.92	-8.29	-3.37	-14.92	-8.76
δd_{23}	-10.0 ± 5.0	+8.53	+9.00	+3.09	+3.99	+3.36
d_{adatom}		2.64	2.74	2.68	2.54	2.59
$E_{1,\text{hop}}$	0.84 ± 0.1	1.10	0.74	0.21	0.23	0.23
$E_{1,\text{exc}}$	0.78 ± 0.1	0.91	0.81	0.50	0.41	0.24
$E_{1,\text{ads}}$		5.42	4.54	5.08	5.44	5.15
$E_{2,\text{ads}}$		5.53	4.66	5.25	5.63	5.41
$E_{2,\text{bind}}$		0.22	0.26	0.33	0.38	0.53
(100)						
δd_{12}	$+0.2 \pm 2.6$	-2.46	-1.41	+0.03	-4.27	-4.25
δd_{23}		-0.80	-0.94	+2.32	+3.54	-0.54
d_{adatom}		2.63	2.70	2.61	2.40	2.47
$E_{1,\text{hop}}$		1.06	0.65	0.65	0.68	0.95
$E_{1,\text{exc}}$	0.47 ± 0.1	0.43	0.41	0.83	0.29	0.36
$E_{1,\text{ads}}$		5.22	4.08	4.75	5.06	4.45
$E_{2,\text{ads}}$		5.31	4.30	4.91	5.22	4.74
$E_{2,\text{bind}}$		0.20	0.13	0.34	0.47	0.56

All energy values are given in units of eV, while d_{adatom} is shown in Å, and δd_{xy} is provided in %. All properties shown here are calculated using $T = 0$ K methods (optimizations or NEB)

mechanisms, but perform well for the exchange profile. For the (110) surface hop and exchange, our ML models are again in excellent agreement with both DFT and experiments.⁶⁸ All EAM potentials significantly underestimate the activation energy of both mechanisms, with only EAM-B predicting a reasonable energy barrier for the exchange mechanism. For the (100) surface hop and exchange, our ML models are once again in excellent agreement with both DFT and experiments.⁶⁷ The various EAM potentials fare slightly better here, though significant discrepancies exist for EAM-B's exchange and EAM-Z's hop predictions.

We finally consider the adsorption and binding energies of a dimer on the (111), (110), and (100) surfaces, which can be found in Table II. The DFT-calculated energies are in good agreement with reported literature values,⁶⁹ where available. Similarly to the adatom adsorption energy, our ML models, while predicting the exact trend in the dimer adsorption energy with respect to DFT, yields values approximately 1 eV lower than DFT. For all surfaces, there exists a spread among EAM predicted values, but again, the trend in dimer

adsorption energy with respect to DFT is captured correctly. Our ML models predict the dimer binding energy, on all surfaces, to be in excellent agreement with the calculated DFT values. EAM-B performs well for all surfaces, while EAM-F deviates from DFT on both the (110) and (100) surfaces, and EAM-Z deviates significantly from DFT for all surfaces considered.

Grain Boundaries

Accurately predicting properties, such as the grain boundary energy, surface energy, and work of separation energy, is an important step in simulating the complex behavior of material defect classes.^{70–75} Recent work has provided a simple prescription of the creation of grain boundary structures, as well as a list of low-*sum* grain boundary, surface, and work of separation energies to benchmark our DFT values against.⁷⁶ Figure 2 provides the reader with a visual representation of all grain boundaries considered in this work, while Fig. 3a encapsulates the data as a parity plot (see Supplementary Table S5 for more details).

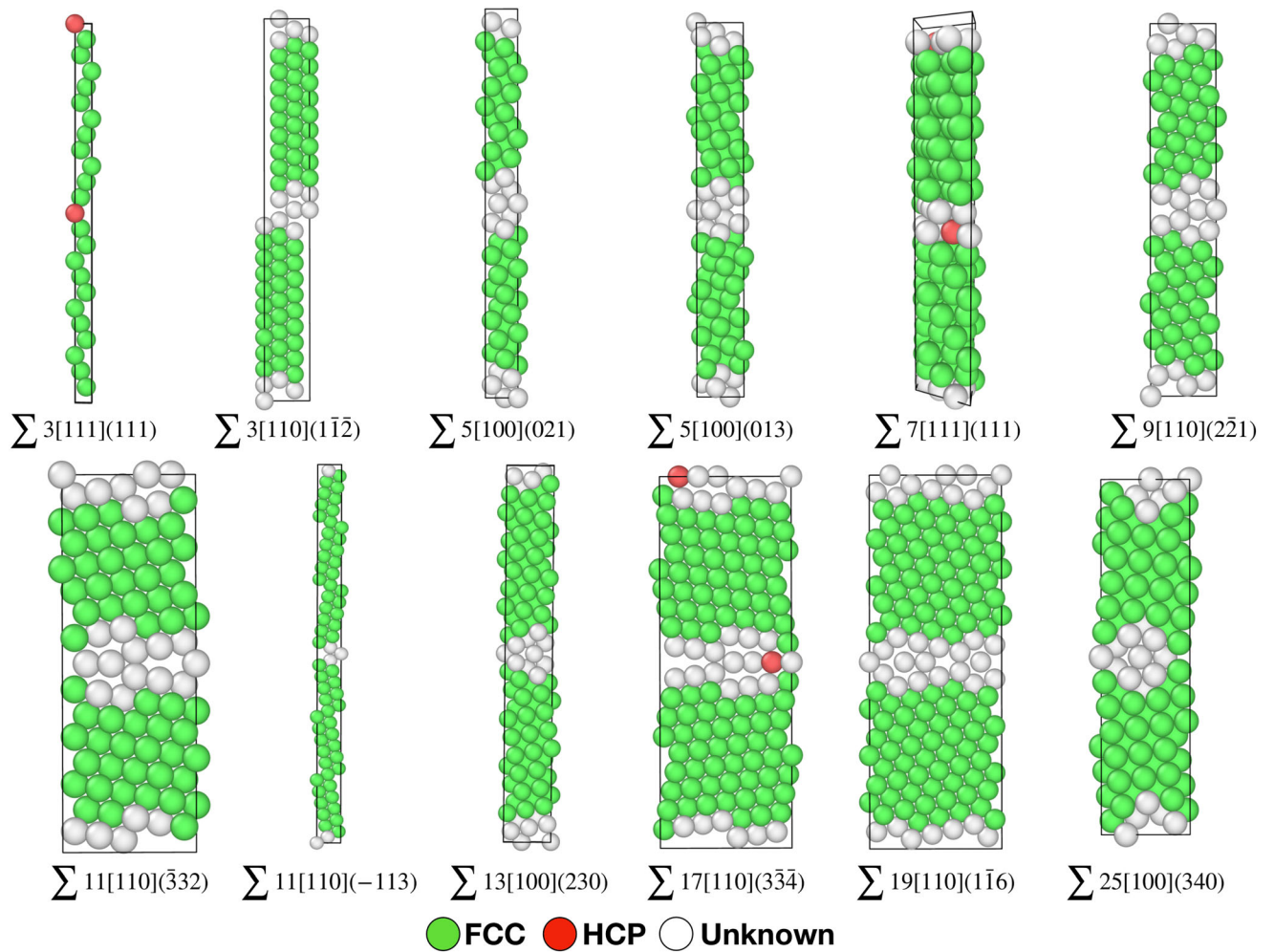


Fig. 2. All grain boundary structures predicted in this work, prior to relaxation. Colors correspond to an atom's coordination environment.

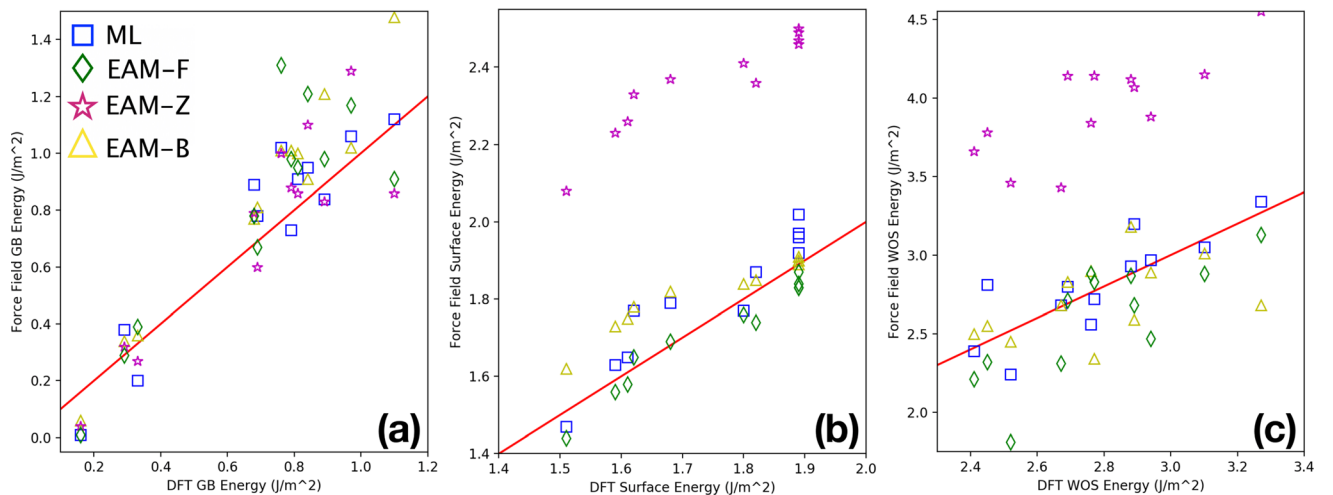


Fig. 3. Grain boundary energy (a), surface energy (b), and work of separation energy (c) calculated for all grain boundaries and surfaces studied in this work. The parity line corresponds to the DFT prediction.

Statistically, our ML models have a root-mean-square error (RMSE) of 0.13 J/m^2 with respect to the DFT grain boundary energies. EAM-B, EAM-Z,

and EAM-F yield RMSE values of 0.19 J/m^2 , 0.17 J/m^2 , and 0.23 J/m^2 , respectively. From these values, the agreement between DFT and our ML models

can clearly be seen. It is also worth noting that the agreement between DFT and ML, compared with the agreement between DFT and EAM, improves as the complexity of the boundary plane increases. This indicates that our ML models could be used to accurately explore the grain boundary energy for boundaries more complex than $\Sigma 25$; a region in which the EAM potentials considered here are likely to deviate from DFT (see Supplementary Table S7 for more details).

The combination of surface and grain boundary energies can be used to predict the work of separation energy, or the energy required to cleave the grain boundary into two free surfaces. Figure 3c shows these data as a parity plot. Statistically, our ML models have an RMSE of 0.17 J/m^2 with respect to the DFT grain boundary energies. EAM-B, EAM-Z, and EAM-F yield RMSE values of 0.26 J/m^2 , 1.17 J/m^2 , and 0.29 J/m^2 , respectively. As our ML models are in good agreement with DFT for both grain boundary and surface energies, the work of separation energies are also predicted extremely well. As all EAM potentials considered in this work perform inadequately for either surface energies or grain boundary energies, the combination of them compounds their errors (see Supplementary Table S5 for more details).

Grain Boundaries with Vacancies

In a dynamic environment, point defects diffuse in and around grain boundaries, making their way from one grain to another.^{77–79} Therefore, understanding these environments at the atomic level allows us to fundamentally understand whether grain boundaries aid or prohibit the diffusion of point defects such as vacancies. To this end we calculated the vacancy formation energy of the grain boundaries discussed in the last few paragraphs. However, the vacancy formation energy was calculated with respect to two distinct reference environments, providing unique insight into how vacancies affect the stability of both grain boundaries, as well as the vacancy diffusion process.

The first reference environment considered was a perfect bulk configuration. By using a bulk configuration as the reference, a direct comparison with the bulk vacancy formation energy can be made, providing insight into the thermodynamic stability of a vacancy within the grain boundary, compared with the bulk-like region of a grain. Using Eq. 10 and following the prescription described earlier in this work, the vacancy formation energy was calculated for all grain boundaries studied. Figure 4a provides the results of these calculations as a parity plot (see Supplementary Table S6 for more details). Statistically, our ML models have an RMSE of 0.26 eV with respect to the DFT grain boundary energies. EAM-B, EAM-Z, and EAM-F yield RMSE values of 1.43 eV , 1.77 eV , and 1.58 eV , respectively. Our ML models clearly show a significant

improvement over the existing EAM potentials, for every considered grain boundary configuration.

It should also be noted that such environments were not included in the ML model's respective training sets. Therefore, we conclude that the ML models can reliably extrapolate to such configuration domains, using precursor information such as pristine grain boundaries, and vacancies in a bulk configuration. Such an ability is crucial when simulating the phenomena occurring in a material containing hundreds of thousands (or more) of atoms, as it is impossible to train the ML models on every possible permutation of atomic configurations during a dynamic process.

The second reference environment considered was a grain boundary configuration without a vacancy. By using the grain boundary as the reference, one can compare the relative thermodynamic stability of the pristine grain boundary versus the boundary in the presence of a vacancy. Here, a positive value indicates that the pristine grain boundary would be more energetically favorable, compared with a negative value, which indicates that the vacancy within the boundary plane is more favorable. Figure 4b provides the results of these calculations as a parity plot (see Supplementary Table S6 for more details). Statistically, our ML models have a root-mean-square error (RMSE) of 0.66 eV with respect to the DFT grain boundary energies. EAM-B, EAM-Z, and EAM-F yield RMSE values of 1.30 eV , 4.45 eV , and 5.64 eV , respectively.

It can be seen from Fig. 4b that many prediction made by the EAM potentials actually indicate that the vacancy within the plane is more energetically stable, a result in stark contrast to DFT. Therefore, while the EAM potentials provide a good understanding of properties for the grain boundaries and surfaces, the introduction of a second defect within these configurations pushes their predictions into pure extrapolation, where their accuracy breaks down. However, our ML models clearly have the ability to make reliable predictions in this regime, and are a clear improvement over the existing models for Pt.

However, the question of how likely it is that diffusion of vacancies along or away from the grain boundary plane will occur is not completely answered by the calculations above. To truly probe the kinetics of the diffusion process itself, NEB calculations were used to study these mechanisms. Here we consider a vacancy diffusing in and around the $\Sigma 3(111)[111]$ grain boundary plane. A $3 \times 3 \times 1$ supercell containing 215 atoms was used (compared with the 27-atom cell used to calculate the grain boundary energy) to avoid the vacancy interacting with its periodic image. Due to the supercell sizes required to achieve this, only $\Sigma 3(111)[111]$ was considered in this work, and we leave a more thorough analysis for future work.

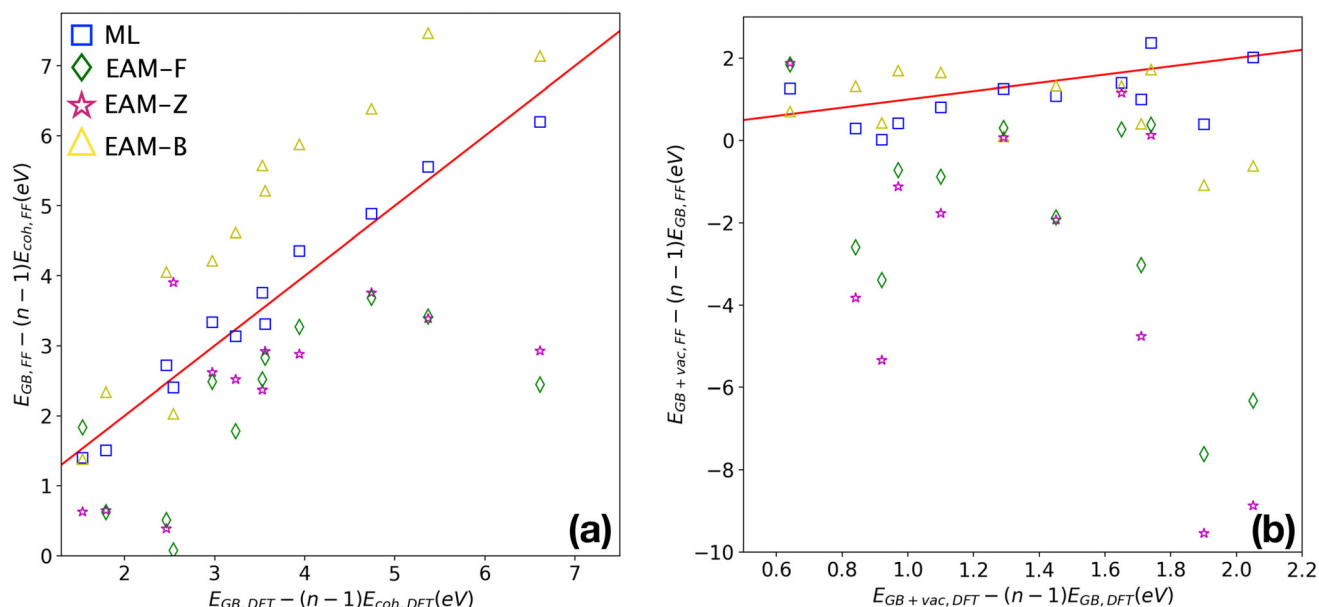


Fig. 4. Vacancy formation energy of single vacancy located inside a grain boundary plane, using (a) bulk and (b) corresponding perfect grain boundary, as the reference system.

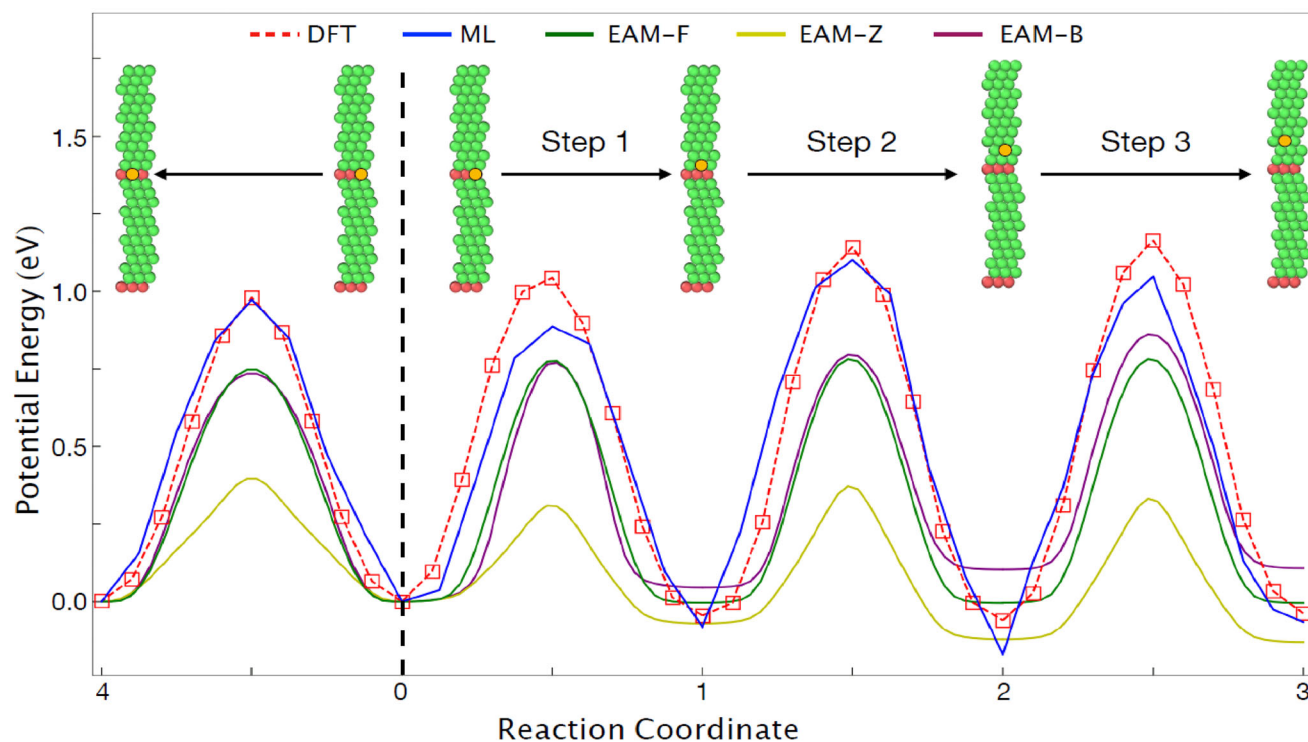


Fig. 5. NEB-calculated diffusion pathways for a monovacancy diffusing in and away from the $\Sigma 3(111)[111]$ grain boundary plane. Reaction coordinates correspond to the pathway taken by the vacancy. The pathway taken along each reaction coordinate is described visually above the energy profiles.

Figure 5 shows the minimum energy profile for the two diffusion pathways: (1) a vacancy moving to an adjacent site within the grain boundary plane, defined as moving from reaction coordinate 0 to 4 in Fig. 5, and (2) a vacancy moving along the grain boundary normal into the bulk portion of the grain,

defined as moving, in order, from reaction coordinate 0 to 1, then 1 to 2, and finally 2 to 3 in Fig. 5. From these calculations, one can see an increase in the DFT transition state energies as the vacancy diffuses away from the grain boundary plane, but also a negative energy difference between the initial

configuration (vacancy in plane) and the final configuration (vacancy 10 Å away from boundary plane), which is consistent with the thermodynamic properties discussed previously. Our ML models follow this trend, lining up with both DFT, but also the ML predictions of the grain boundary vacancy formation energy.

However, the EAM potential predictions disagree with not only DFT but also with each other. All three EAM potentials predict a unique relationship between the initial and final configurations, for the pathway moving the vacancy away from the boundary plane. Only EAM-Z predicts the correct qualitative relationship, in that the final configuration is lower in energy than the initial configuration. However, this comes at the cost of transition state energies being nearly four times lower than that of DFT, indicating that the diffusion mechanism occurs more frequently at lower temperatures. EAM-B and EAM-F, while predicting more accurate transition state energies, indicate that the vacancy is either equally likely, or more likely, to end up staying within the boundary plane compared with migrating into the bulk-like region. This is inconsistent with both the DFT-NEB prediction and also the respective EAM calculations of the grain boundary vacancy formation energy, indicating that these EAM potentials are not versatile enough to study these configuration spaces.

Grain Coarsening

While the previous section aimed to capture an atomic-level understanding of different defect classes, and their interactions with each other, larger length and time scales must be explored to truly connect with experimental observations. To this end, we consider the phenomena of grain coarsening, as a function of temperature, through MD (*NPT*) simulations, in which the ML stress model is used to drive changes in the cell volume. Here, we aim to connect to experimental observations of annealing nanocrystalline platinum during an irreversible recovery process⁴² by calculating the average grain size as a function of temperature.

Experimentally, below 175°C, there is no change in the average grain size observed, outside of statistical fluctuations. It has been proposed that this behavior is due to the relaxation of unstable grain boundaries to their respective metastable configurations.⁴² Between 175°C and 200°C, the mean grain size begins to increase, and continues to increase through 325°C.

Figure 6 (top) shows the distribution of grain sizes, along with a fit, normalized Gaussian function, for several temperatures around the transition point between metastable grain-boundary transition, and grain growth. Figure 6 (bottom) shows the mean grain sizes plotted as a function of each temperature studied in this work (see Supplementary Fig. S2 for more details). The *NPT* simulations

performed using our ML models predict a transition point at exactly the same temperature as experiments, showing excellent agreement between ML and experiments. We also extend the temperature range beyond that of experiments, to 425°C. While our MD simulations indicate that grain growth plateaus between 325°C and 350°C, such a feature could be an artifact of the system size, and larger system sizes may not yield such behavior.

There are two critical pieces of information to note from these MD simulations. First, we emphasize here that, with the system sizes considered in this work, we cannot attain mean grain sizes for direct comparison with experiments. Indeed, previous MD work on grain growth has indicated that system sizes containing more than 16 million atoms would be needed to quantitatively compare the grain structure with experiments. The aim of this work is not to predict the exact grain structure, but rather to determine the threshold temperature for grain growth.

CONCLUSION

A variety of defect classes are studied, over length scales ranging from a few angstroms to tens of nanometers, using machine learning models of atomic forces, potential energy, and stress, trained on DFT data. Thermodynamic properties of grain boundaries, such as the grain boundary energy, work of separation energy, and vacancy formation energy of a vacancy within the grain boundary, were calculated for 12 grain boundaries. The minimum-energy pathway for vacancy diffusion along, and away from, the $\Sigma 3(111)[111]$ grain boundary was also studied. Surface energy was calculated for the surfaces corresponding to each grain boundary. A more detailed analysis of surface phenomena on Pt (111), (110), and (100), involving adatom and dimer adsorption energy, dimer binding energy, and activation energy of a single adatom either hopping on the surface, or exchanging places with a surface atom, was also performed. Finally, the growth of grains was predicted via MD simulations, employing systems containing over half a million atoms. When possible, these results were compared with DFT and EAM calculations, or directly with experiments.

The ML models presented in this work show a clear improvement over current EAM potentials, when considering the predictions made over all of the various tests performed in this work. The ML models used here also represent a paradigm shift, in that they were not constructed from scratch to calculate the properties studied in this work. Rather, these ML models were iteratively improved, using a previous ML model as a starting point. They also do not contain reference data for all of the environments studied here, indicating their extrapolative power. Using ML models, with the accuracy of DFT and the speed of classical/semiempirical

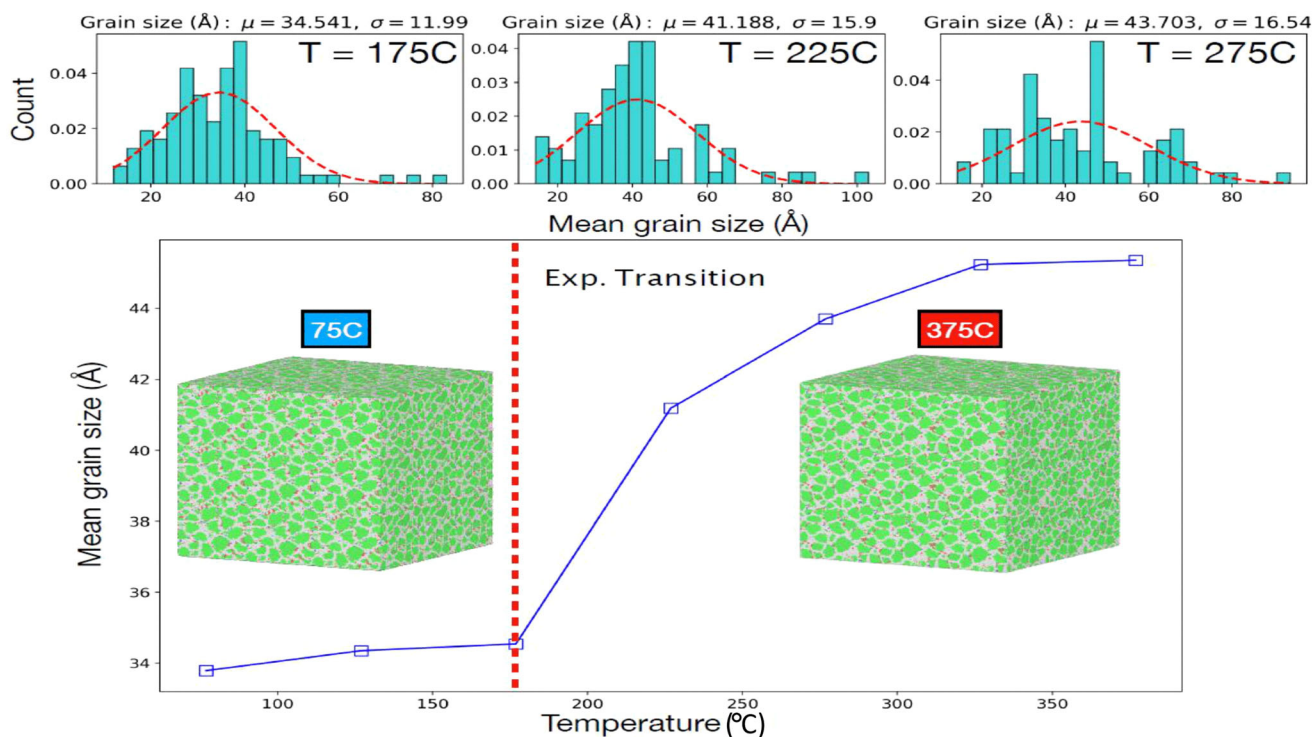


Fig. 6. (Top) Normalized histograms detailing the grain sizes calculated from MD simulations at temperatures between 175°C and 275°C. Mean grain size and spread of grain sizes can be seen above each histogram. (Bottom) Predicted grain size as a function of temperature. The vertical dotted line corresponds to the experimental transition temperature, at which grain grows begins. The resulting grain structure from the ML MD simulations for 75°C and 375°C are shown to the left and right, respectively.

models, allows us to accurately and reliably study properties that are impractical to compute using DFT. This work adds another layer of validation that ML models can make reliable predictions over multiple length and time scales, and solidifies ML as a vital tool in the study of atomic and nanoscale research.

ACKNOWLEDGEMENTS

This work was supported financially by the National Science Foundation (Grant No. 1821992) and the Office of Naval Research (Grant No. N000141812113). The authors would also like to thank Dr. Ghanshyam Pilia for his helpful insight regarding dislocation configurations.

DATA AVAILABILITY

All data used to reproduce the results of this work can be found at khazana.gatech.edu.

CONFLICT OF INTEREST

On behalf of all authors, the corresponding author states that there are no conflicts of interest.

ELECTRONIC SUPPLEMENTARY MATERIAL

The online version of this article (<https://doi.org/10.1007/s11837-020-04385-0>) contains supplementary material, which is available to authorized users.

REFERENCES

1. V. Botu, J. Chapman, and R. Ramprasad, *Comput. Mater. Sci.* 129, 332 (2016).
2. S.M. Rassoulinejad-Mousavi and Y. Zhang, *Sci. Rep.* 8, 2424 (2018).
3. G. Gregory, S.P. Russo, and I.K. Snook, *J. Chem. Phys.* 123, 204719 (2005).
4. I.J. O'Brien, C.M. Barr, P.M. Price, K. Hattar, and S.M. Foiles, *J. Mater. Sci.* 53, 2911 (2018).
5. X.W. Zhou, R.A. Johnson, and H.N.G. Wadley, *Phys. Rev. B* 60, 144113 (2004).
6. S.M. Foiles, M.I. Baskes, and M.S. Daw, *Phys. Rev. B* 59, 11693 (1999).
7. J. Chapman, R. Batra, B.P. Uberuaga, G. Pilia, and R. Ramprasad, *Comput. Mater. Sci.* 158, 353 (2019).
8. P. Hohenberg and W. Kohn, *Phys. Rev.* 126, B864 (1964).
9. W. Kohn and L. Sham, *Phys. Rev.* 140, A1133 (1965).
10. R.O. Jones, *Rev. Mod. Phys.* 87, 897 (2015).
11. J.E. Jones and S. Chapman, *Proc. R. Soc. Lond. A* 106, 463 (1924).

12. M.S. Daw and M.I. Baskes, *Phys. Rev. B* 29, 6443 (1984).
13. M.S. Daw, S.M. Foiles, and M.I. Baskes, *Mater. Sci. Rep.* 9, 251 (1993).
14. J. Tersoff, *Phys. Rev. B* 37, 6991 (1998).
15. M.Z. Bazant, E. Kaxiras, and J.F. Justo, *Phys. Rev. B* 56, 8542 (1997).
16. A.C.T. van Duin, S. Dasgupta, F. Lorant, and W.A. Goddard, *J. Phys. Chem. A* 105, 9396 (2001).
17. A.R. Setoodeh and H. Attariani, *Mater. Lett.* 62, 4266 (2008).
18. D. Poulidakos, S. Arcidiacono, and S. Maruyama, *Micro. Thermophys. Eng.* 181 (2003).
19. B. Baretzky, M.D. Baró, G.P. Grabovetskaya, J. Gubicza, M.B. Ivanov, T.G. YuR Kolobov, J. Langdon, A.G. Lendvai, A.A. Lipnitskii, A.A. Mazilkin, J. Nazarov, I.A. Nogués, S.G. Ovidko, G.I. Protasova, A. Raab, N.V. Révész, M.J. Skiba, B.B. Starink, S. Straumal, T. Suriñach, and A.P.Z. Ungár, *Rev. Adv. Mater. Sci.* 9, 45 (2005).
20. J. Gasteiger and J. Zupan, *Angew. Chem. Int. Ed.* 32, 503 (1993).
21. B.G. Sumpter, C. Getino, and D.W. Noid, *Annu. Rev. Phys. Chem.* 45, 439 (1994).
22. K. Rajan, *Mater. Today* 8, 38 (2005).
23. R. Ramprasad, R. Batra, G. Pilania, A. Mannodi-Kanakkithodi, and C. Kim, *NPJ Comput. Mater.* 54 (2017).
24. A. Mannodi-Kanakkithodi, A. Chandrasekaran, C. Kim, T.D. Huan, G. Pilania, V. Botu, and R. Ramprasad, *Mater. Today* 21, 785 (2017).
25. C. Kim, A. Chandrasekaran, T.D. Huan, D. Das, and R. Ramprasad, *J. Phys. Chem. C* 122, 17575 (2018).
26. G. Pilania, C. Wang, X. Jiang, S. Rajasekaran, and R. Ramprasad, *Sci. Rep.* 3, 2810 (2018).
27. T.D. Huan, A. Mannodi-Kanakkithodi, and R. Ramprasad, *Phys. Rev. B* 92 (2015).
28. A. Mannodi-Kanakkithodi, G. Pilania, T.D. Huan, T. Lookman, and R. Ramprasad, *Sci. Rep.* 6, 20952 (2016).
29. V. Botu and R. Ramprasad, *Int. J. Quant. Chem.* 115, 1074 (2015).
30. V. Botu and R. Ramprasad, *Phys. Rev. B* 92, 094306 (2015).
31. V. Botu, R. Batra, J. Chapman, and R. Ramprasad, *J. Phys. Chem. C* 121, 511 (2017).
32. T.D. Huan, R. Batra, J. Chapman, S. Krishnan, L. Chen, and R. Ramprasad, *NPJ Comput. Mater.* 3, 37 (2017).
33. J. Behler and M. Parrinello, *Phys. Rev. Lett.* 98, 146401 (2007).
34. A.P. Bartók and G. Csányi, *Int. J. Quant. Chem.* 115, 1051 (2015).
35. S. Chmiela, A. Tkatchenko, H.E. Sauceda, T. Poltavsky, K.T. Schutt, and K.R. Muller, *Sci. Adv.* 3, e1603015 (2017).
36. T.D. Huan, R. Batra, J. Chapman, C. Kim, A. Chandrasekaran, and R. Ramprasad, *J. Phys. Chem. C* 123, 20715 (2019).
37. J. Chapman, R. Batra, and R. Ramprasad, *Comput. Mater. Sci.* 174, 109483 (2020).
38. V.L. Deringer and G. Csányi, *Phys. Rev. B* 95, 094203 (2017).
39. W.J. Szlachta, A.P. Bartók, and G. Csányi, *Phys. Rev. B* 90, 104108 (2014).
40. J. Byggmästar, A. Hamedani, K. Nordlund, and F. Djurabekova, *Phys. Rev. B* 10, 144105 (2019).
41. J. Chapman and R. Ramprasad, *J. Chem. Phys.* 152, 224709 (2020).
42. A. Tschöpe and R. Birringer, *Acta Metall. Mater.* 41, 2791 (1993).
43. R. Batra, T.D. Huan, C. Kim, J. Chapman, L. Chen, A. Chandrasekaran, and R. Ramprasad, *J. Phys. Chem. C* 123, 15859 (2019).
44. A. Chandrasekaran, D. Kamal, R. Batra, C. Kim, L. Chen, and R. Ramprasad, *NPJ Comput. Mater.* 5, 22 (2019).
45. G. Kresse and J. Furthmüller, *Phys. Rev. B* 54, 11169 (1996).
46. G. Kresse and D. Joubert, *Phys. Rev. B* 59, 1758 (1999).
47. H. Jönsson, G. Mills, and K.W. Jacobsen, *Classical and Quantum Dynamics in Condensed Phase Simulations* (World Scientific, 1998), p. 385.
48. H. Jönsson and G. Henkelman, *J. Chem. Phys.* 113, 9978 (2000).
49. H. Jönsson, G. Henkelman, and B.P. Uberuaga, *J. Chem. Phys.* 113, 9901 (2000).
50. J.P. Perdew, K. Burke, and Y. Wang, *Phys. Rev. B* 54, 16533 (1996).
51. P.E. Blöchl, *Phys. Rev. B* 50, 17953 (1994).
52. H.J. Monkhorst and J.D. Pack, *Phys. Rev. B* 13, 5188 (1976).
53. S. Plimpton, *J. Comput. Phys.* 117, 1 (1995).
54. J. Cheng, J. Luo, and K. Yang, *Comput. Mater. Sci.* 155, 92 (2018).
55. P. Hirel, *Comput. Phys. Commun.* 197, 212 (2015).
56. J. Li, B. Lu, H. Zhou, C. Tian, Y. Xian, G. Hu, and R. Xia, *Phys. Lett. A* 383, 1922 (2019).
57. Y. Shibuta, K. Oguchi, T. Takaki, and M. Ohno, *Sci. Rep.* 5, 13534 (2015).
58. A. Hunter, R.F. Zhang, and I.J. Beyerlein, *J. Appl. Phys.* 15, 134314 (2014).
59. F.J. Zerilli and R.W. Armstrong, *Acta Metall. Mater.* 40, 1803 (1992).
60. S. Mendelson, *J. App. Phys.* 41, 1893 (1970).
61. H. Van Swygenhoven and J. Weertman, *Mater. Today* 9, 24 (2006).
62. Z.H. Jin, S.T. Dunham, H. Gleiter, H. Hahn, and P. Gumbsch, *Scr. Mater.* 64, 605 (2011).
63. A. Stukowski and K. Albe, *Modell. Simul. Mater. Sci. Eng.* 18, 085001 (2010).
64. G. Boisvert, L.J. Lewis, and M. Scheffler, *Phys. Rev. B* 57, 1881 (1998).
65. N.E. Singh-Miller and N. Marzari, *Phys. Rev. B* 80, 235407 (2009).
66. N. Panagiotides and N.I. Papanicolaou, *Int. J. Quant. Chem.* 110, 202 (2010).
67. M. Villarba and H. Jönsson, *Surf. Sci.* 317, 15 (1994).
68. D.W. Bassett and P.R. Webber, *Surf. Sci.* 70, 520 (1978).
69. I. Matrone, M. Mazroui, and K. Shibaai, *Chem. Phys.* 502, 96 (2018).
70. G.S. Rohrer, *J. Mater. Sci.* 46, 5881 (2011).
71. A.D. Rollett, G. Gottstein, L.S. Shvindlerman, and D.A. Molodov, *Z. Metall.* 95, 226 (2004).
72. R.C. Gifkins, *Mater. Charact.* 32, 59 (1994).
73. L. Zhang, C. Lu, and K. Tieu, *Comput. Mater. Sci.* 118, 180 (2016).
74. R.W. Balluffi, Y. Komem, and T. Schober, *Surf. Sci.* 31, 68 (1972).
75. G. Reiss, J. Vancea, and H. Hoffmann, *Phys. Rev. Lett.* 56, 2100 (1986).
76. H. Zheng, X. Li, R. Tran, C. Chen, M. Horton, D. Winston, K. Persson, and S.P. Ong, *Acta Mater.* 186, 40 (2020).
77. M.R. Sørensen, Y. Mishin, and A.F. Voter, *Phys. Rev. B* 62, 3658 (2000).
78. M.A. Tschopp, M.F. Horstemeyer, F. Gao, X. Sun, and M. Khaleel, *Scr. Mater.* 64, 908 (2011).
79. M. Upmanyu, D.J. Srolovitz, L.S. Shvindlerman, and G. Gottstein, *Interface Sci.* 6, 289 (1998).

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.