# Examination of computed aluminum grain boundary structures and energies that span the 5D space of crystallographic character

Eric R. Homer[a,*], Gus L. W. Hart[b], C. Braxton Owens[b], Derek Hensley[b], Jay Spendlove[b], Lydia Harris Serafin[b]

[a]*Department of Mechanical Engineering, Brigham Young University, Provo, UT, 84602, USA.*
[b]*Department of Physics and Astronomy, Brigham Young University, Provo, UT, 84602, USA.*

**Abstract**

The space of possible grain boundary structures is vast, with 5 macroscopic, crystallographic degrees of freedom that define the character of a grain boundary. While numerous datasets of grain boundaries have examined this space in part or in full, we present a computed dataset of 7304 unique aluminum grain boundaries in the 5D crystallographic space. Our sampling also includes a range of possible microscopic, atomic configurations for each unique 5D crystallographic structure, which total over 43 million structures. We present the methods used to generate this dataset, an initial examination of the energy trends that follow the Read-Shockley relationship, hints at trends throughout the 5D space, variations in GB energy when non-minimum energy structures are examined, and insights gained in machine learning of grain boundary energy structure-property relationships. This dataset, which is available for download, has great potential for insight into GB structure-property relationships.

*Keywords:* Grain boundaries, Atomistic simulations, Aluminum

## 1. Introduction

The interfaces or grain boundaries (GBs) separating the individual crystals in polycrystalline materials can have a significant effect on numerous material properties. In fact, in the 1990s, a field emerged called GB engineering that demonstrated large improvements in properties by controlling the population of different GB types [1–3]. Notable achievements in GBE include a decreased intergranular corrosion susceptibility [4], a gain in cycle life of lead acid batteries [5], a decrease in creep rate [6], an increased strength, ductility, fracture toughness [7], among others. Unfortunately, many of these improvements were limited to low stacking fault energy FCC materials that readily form twins. To enable GBE generally we need more information about the range of energies and structures that are possible in different GB types. In other words, we need to understand GB structure-property relationships over the full range of GB types.

GB structures are defined by both their macroscopic and microscopic degrees of freedom. The macroscopic degrees of freedom are defined by five crystallographic parameters, often referred to as the GB character. Three parameters describe the misorientation between two grains and two parameters define the orientation of the boundary plane (BP) separating the two grains. The microscopic degrees of freedom are defined at the most basic level by the positions of each atom, where $N$ atoms would have $3N$ degrees of freedom. Various metrics have been developed to describe the atomic structure of GBs in simpler terms, including the structural unit model [8–10], polyhedral unit model [11], and more recently by various collections of local atomic environments [12–22].

It must also be noted that for every unique macroscopic structure description, there are a multiplicity of microscopic degrees of freedom where the atoms are arranged differently [14, 23–27]. These various configurations are typically referred to as metastable configurations or GB phases. In experiments, these can be seen as varying atomic configurations along a single GB [27]. In modeling and simulation, most research focuses on the minimum energy configurations. However, a statistical mechanical approach using these metastable states can be used to predict finite temperature equilibrium and non-equilibrium properties [25].

While the 5 macroscopic degrees of freedom are significantly fewer than the $3N$ microscopic degrees of freedom, especially considering the metastable states, they still constitute a large space that is challenging to fully resolve. Symmetries are present in this 5D space that reduce the amount of data required to explore it [28], but these symmetries can be as problematic as they are helpful because researchers do not always use the unique descriptions of GBs, such as the disorientation[1] or a unique definition of the BP, such as described in [29].

Historically, these 5 degrees of freedom have frequently been simplified down to one degree of freedom, completely ignoring BP and disorientation axis, focusing only on the disorientation angle. This is often simplified even further into low vs. high disorientation angles or special vs. not special, where special GBs

*Corresponding author
*Email address:* eric.homer@byu.edu (Eric R. Homer)

[1] A disorientation provides a single unique description for many symmetrically equivalent misorientations and is defined by the minimum misorientation angle when the misorientation axis is found in the standard stereographic triangle.

are comprised of low Σ-value coincidence site lattices (CSLs) and low disorientation angle boundaries. While these simplifications have their uses, they represent an incomplete characterization of the GB.

There have been many notable investigations of GB structure-property relationships over numerous different properties, but it is beyond the scope of this work to provide a full review of these efforts beyond highlighting some important works [30–34]. However, we do wish to highlight those studies that have either been particularly insightful or unique in their investigation of GB interface energy over some portion, if not all, of the 5D macroscopic GB space. The most well-known of these relationships is the Read-Shockley relationship, which predicts GB energy and atomic structure as a function of disorientation angle, in the most basic case for tilt or symmetric tilt GBs [35]. This relationship has proven remarkably useful over the years and will see reinforcement in this work. Other noteworthy models include the broken bond GB energy model by Wolf [36, 37], the Frank-Bilby equation to relate microscopic and macroscopic degrees of freedom [38–40] and the Wulff construction and Cahn-Hoffman capillarity vector to understand the role of BP [41–44]. More recently Bulatov, Reed, and Kumar created a GB energy function for FCC materials that predicts energy for four different FCC metals based upon the 5 macroscopic degrees of freedom [45].

Of course all these models require data upon which to learn and predict. Numerous model, experimental, and computed datasets have been developed to provide insight into GB energy structure-property relationships. Older datasets relied on hard sphere approximations [46] while newer datasets take advantage of advances in characterization and computation. For example, advances in computing have also allowed larger and more detailed atomic structures of GBs to be simulated. Recent experimental datasets take advantage of automated serial sectioning technologies [47–49] to recover all 5 macroscopic degrees of freedom and morphology of polycrystalline microstructures.

While datasets of computed GBs date back several decades [50–53], a dataset created by Olmsted, Foiles, and Holm [54] has received particular attention in the literature. The dataset consists of 388 GBs of Ni and Al and their corresponding minimum energies and atomic structure configurations. The 388 GBs are comprised of 72 unique disorientations where some disorientations sample many BPs, such as the Σ3, while others sample as few as 2 BPs. Numerous other computed GB datasets have since been created including 408 GBs in BCC Fe and Mo [55], 174 symmetric and asymmetric tilt GBs in both Al and Cu [56], 126 [1 0 0] symmetric tilt Ni GBs [18, 57], 346 [1 0 0] disorientation axis Ni GBs [57, 58], 344 GBs in Si and SiC [59], 230 GBs in both $UO_2$ and $CeO_2$ [60], more than 68 000 GBs in Fe [61], and 327 GBs spanning 58 different elements using density functional theory [62]. Additional detail about each of these computational datasets is provided in the supplemental material. The dataset from Kim et al. is by far the largest but it should be noted that each GB only examined a single initial condition [61], so it is not clear that this is representative of the minimum GB energy typical to most datasets.

Experimental datasets of GBs that recover all 5 macroscopic degrees of freedom often infer GB energy values based on population, geometry of the GB network, or thermal grooving. For example, one can utilize the Herring equation or Cahn-Hoffman capillarity vector to reconstruct the GB energy based on the geometry of triple junctions, where three GBs come together and an energy balance of the three GBs would allow the triple junction to achieve equilibrium [33, 63, 64]. One can also infer GB energy based on GB populations, where it is assumed that low energy GBs would show up with high frequency and high energy GBs would show up with low frequency [65]. Using one of these methods, GB energy has been, or can be, inferred in the following experimentally obtained GB datasets. These include measurements in Ni [66], Al [67, 68], NiAl [69], ferritic steel [70, 71], GB engineered Ni and Cu [72], yttria [48], and magnesia [65], among others [73], of which many of these datasets are available for download [74].

In direct comparisons, Amouyal et al. find good agreement between measured and computed GB energies in NiAl [69]. Barmak et al. find more variation in a limited comparison between measured and computed GB energies in Al [75]. In a large-scale comparison of Ni datasets, Rohrer et al. concluded that the experimental and computational results validate each other for boundaries that are appropriately represented in both data sets [76]. In later work, Holm et al. demonstrate that one can also validate GB energy calculations using the GB character distribution [68].

GB datasets are also crucial to the Materials Genome Initiative focus to develop new materials on a shorter time line, because digital data is identified as one of the three pillars in the materials innovation infrastructure, along with computational and experimental tools [77]. As will be illustrated below, the Olmsted GB dataset is an excellent example of how digital data has been used in numerous studies to examine a variety of phenomena [54].

The FCC GB energy function mentioned earlier, created by Bulatov et al., was trained on the Olmsted dataset and showed that four FCC metals all had similar GB energy trends across the 5D space [45]. Furthermore, the energy prediction could be changed between all four FCC metals by simply changing two parameters related to the twin and random GB energy values.

Olmsted's dataset was also used to gain insight into energy variation through the 5D space. Bulatov et al. found interesting properties about the energy cusps [45]. Homer and Patala built upon Patala's work on the symmetries of BP orientations [78] to identify fundamental zones (or irreducible spaces) in the BP degrees of freedom [29]. Using these BP fundamental zones, they found that GB energy in the Olmsted dataset varied smoothly for similar BP orientations of the same disorientation as a function of disorientation angle. This trend was later confirmed more generally by Erickson and Homer in their dataset of 346 [1 0 0] disorientation axis GBs [58].

Olmsted's dataset has also been used to examine GB mobility [79, 80], GB shear coupling [81, 82], excess volume [83], and others [19, 84, 85]. More recently Olmsted's dataset of atomic structures were used as input to predict GB energy by machine learning [15, 17, 21]. Priedeman et al. built on those meth-

ods and used dimensionality reduction techniques to illustrate a connection between the macroscopic and microscopic degrees of freedom [18].

The other GB datasets have been used for numerous purposes. For example, Zheng developed an improved predictive model for the GB energy of different elements based on the cohesive energy and shear modulus using the DFT dataset mentioned previously [62]. Tamura et al. also developed a predictive model using machine learning, where accurate predictions are reported from training sets with as few as 10 GBs [12]. Restrepo et al. trained an artificial neural network on the Kim dataset to predict GB energies [61, 86].

It is clear that GB datasets have tremendous value, and that, to continue the efforts of GB engineering, more complete structure-property relationships are required. But, to obtain more complete relationships the datasets must span the full range of macroscopic degrees of freedom. Additionally, it would be beneficial if the datasets also provide insight into the range of microscopic degrees of freedom.

In this work, we present a computed dataset of aluminum GBs that span the entire 5D space and provide insight into the range of metastable atomic configurations. In total, 7304 unique aluminum GB structures are examined, with a full set of metastable atomic structure configurations for each. All together, over 43 million GB structures are included in this dataset. In this work we describe the methods used to construct the dataset, examine general statistics of the dataset, illustrate GB energy trends in subspaces of the 5D GB character, illustrate trends in energy distributions of the metastable configurations, and examine basic machine learning predictions using the atomic structures as input. Finally, the dataset will be made available so that it can serve as a tool of understanding to the entire scientific community [87].

## 2. Methods

### *2.1. GB selection procedure*

As noted above, the macroscopic (crystallographic) character of a GB is defined by 5 parameters, 3 for the disorientation and 2 for the boundary plane. In generating our dataset, we selected disorientations that provided comprehensive coverage of the disorientation space, though we did not attempt to use methods for uniform coverage, such as those described Quey et al. [88].

To avoid the effects of free surfaces, periodic boundary conditions are used in the plane of the GB. This is readily solved when using CSL disorientations because the GB plane will be periodic for both grains if CSL lattice points are used when defining the supercell of the simulation. As such, all GBs used in this work are CSLs, but it should be noted that while low $\Sigma$ CSLs are often considered to be special, numerous works contradict this assertion [58, 89–94]. In other words, the CSL framework is used for its utility, not any attempt to sample "special" GBs.

Various factors were considered in selecting the disorientations (CSLs) for the dataset, with the most important being

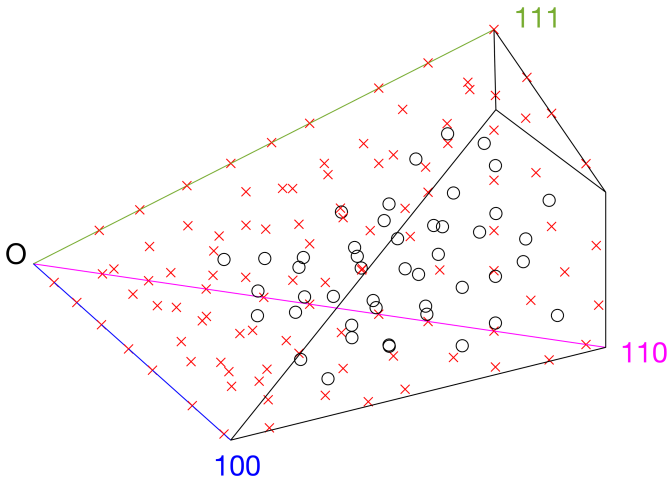


Figure 1: Plot of the 150 selected CSLs in the cubic-cubic fundamental zone in Rodriguez-Frank space. Points on the surface of the fundamental zone are marked with a red 'x' while points on the interior are marked with a black 'o'.

Table 1: Statistics of the selected GBs and BPs, organized by the point group symmetry of the CSL lattice.

| Point Group | Hemisphere Coverage | # CSLs | Average # of BPs/CSL | Total # of BPs |
|---|---|---|---|---|
| $C_i$ | 1/1 | 46 | 54 | 2469 |
| $C_{2h}$ | 1/2 | 75 | 51 | 3834 |
| $D_{2h}$ | 1/4 | 13 | 40 | 525 |
| $D_{3d}$ | 1/6 | 8 | 30 | 240 |
| $D_{4h}$ | 1/8 | 7 | 32 | 223 |
| $D_{6h}$ | 1/12 | 1 | 13 | 13 |
| | Total | 150 | | 7304 |

comprehensive coverage of the 3D disorientation space at $\sim 5°$ spacing. Using the selection procedure described in the supplemental material, we obtained 150 CSLs. Of the CSLs in this set, 46 lie in the interior of the fundamental zone, 7 have [100] disorientation axes, and there are 9 on both the [110] and [111] disorientation axes. These 150 CSLs are plotted in the cubic-cubic disorientation fundamental zone in Rodriguez-Frank space in Figure 1 and listed in supplemental Table S1. It is worth noting that there are 75 CSLs with $\Sigma \leq 99$, 35 CSLs with $99 < \Sigma \leq 199$, 31 GBs with $199 < \Sigma \leq 499$, and 9 GBs with $499 < \Sigma \leq 999$; the largest of which is a $\Sigma$999ae.

From these 150 CSLs, we wanted to select GBs that would provide comprehensive coverage of the unique portions of the BP space for each CSL. The BP symmetries [78] enable the identification of the irreducible space or fundamental zone of unique BPs. The 3D point group symmetry of each CSL is listed in supplemental Table S1. Table 1 provides a summary of all the point group symmetries in this dataset along with the hemisphere coverage of the fundamental zone for that symmetry.

BPs were selected from each fundamental zone, with the goal of providing comprehensive coverage, such that one can interpolate energies over regions that are not sampled. This approach led to much coarser spacing of points than occurred in

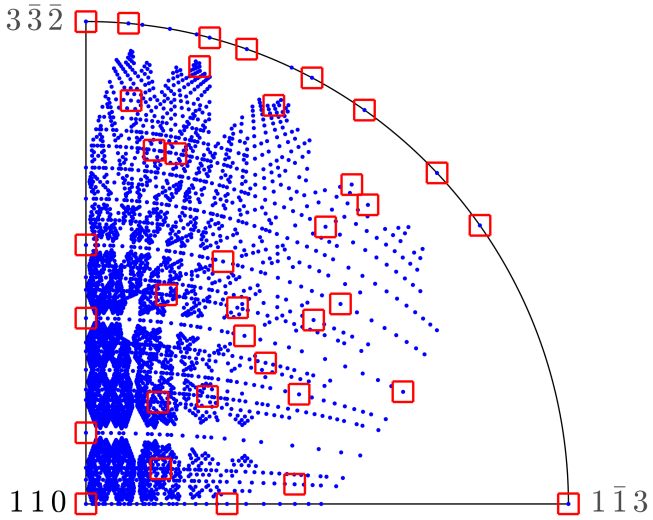


Figure 2: BP fundamental zone plot for the Σ9 CSL marking possible BP orientations in blue that can be selected from ±4 CSL lattice points and marking the selected BP orientations in red squares that aim to provide comprehensive coverage of the BP fundamental zone.

the disorientation fundamental zone, but was necessary to keep the total number of GBs at a reasonable number.

Unfortunately, the method used to select possible BPs does not lead to uniform coverage of each fundamental zone. This results from the use of a CSL lattice to select possible BPs, which automatically satisfy periodic boundary conditions [95]. Additional details on the BP selection procedure are provided in the supplemental materials. Fig. 2 illustrates the possible BPs marked by blue dots that result from this procedure for the Σ9 CSL. It is clear that some BPs are difficult to access using this approach. From these available BPs, we employ an automatic selection procedure for overall coverage. This is followed by a manual selection procedure to fill in any large gaps over the BP fundamental zone. The final BPs selected for the dataset in the Σ9 CSL are marked with red squares in Fig. 2.

A similar selection process of BPs was carried out for all 150 CSLs. The average number of BPs for each CSL is given in Table 1 according to the point group symmetry; it is noted that those numbers do not provide consistent coverage in terms of the number of BPs per unit area of the BP fundamental zone. In the end, a total of 7304 GBs with unique disorientation and BP were selected for the dataset. It is worth noting that this is an order of magnitude greater in size than any other *minimum GB energy* dataset mentioned in the introduction (the dataset of 68 000 GBs created by Kim et al. are not necessarily representative of minimum energy GBs [61]).

### 2.2. Atomic GB structure creation

Once selected, the vectors defining the basis for the periodic GB supercell are used to construct bicrystals. While many GB simulations require the basis vectors in the GB plane of the supercell to be orthogonal, we do not set this as a requirement because it allows a much larger population of BPs to be investigated [95]. We employ monoclinic simulation supercells that contain a single GB. Periodic boundary conditions are used in the GB plane but not in direction normal to the GB. Rigid blocks of atoms at the two extremes normal to the GB mitigate the effects of the free surfaces [54, 96].

In this work we simulate aluminum GB structures using the aluminum EAM potential by Mishin et al. [97]. However, the framework presented here is fairly general and could be applied to other FCC metals or even translated for other pure crystal structures.

To find the minimum energy structures we follow the methods of Olmsted et al. [54] and sample various initial configurations of atom positions. These are obtained by considering the variation of the following six variables: translation of one grain relative to the other within the DSC lattice (3 variables), placement of the boundary interface (1 variable), allowed proximity of atoms before one is selected for deletion (1 variable), and three possible procedures for deleting overlapping atoms (1 variable).

Each of the structures is minimized using conjugate gradient minimization in LAMMPS [98]. On average 5888 structures were minimized for each GB to obtain the minimum energy structure. In total, for all 7304 unique GBs, 43 009 236 total structures have been minimized. After minimization, the GB energy is calculated following standard methods [96]. Note that structures are not thermalized in an attempt to find even lower minimum energy structures; the motivation behind not doing this is that thermalization often leads to faceting of the GBs, providing information about how that particular GB would like to evolve, but not providing a measure of energy for that particular BP orientation. One can then use these energies in a Wulff construction or Cahn-Hoffman capillarity vector [41–44] to determine what the equilibrium structure and faceting would then be from discrete GB measurements, as in [99, 100].

The dataset analyzed in this work focuses mostly on the minimum energy GB structure and corresponding properties of the 7304 unique GBs. These minimum energy GB structures are also analyzed for their GB width, based on the extent of atoms with non-FCC structure, as defined by common neighbor analysis (CNA), as well as for their excess volume per unit GB area, calculated according to standard methods [96], using Python scripts in OVITO [101]. Data relevant to the GB disorientation, BP orientation, crystal orientations, minimum GB energy, number of structures minimized, structure file name, and excess volume, as well as all the minimum energy GB structures, are published in a dataset on Mendeley Data [87].

### 2.3. Machine learning methods

Characterizing microscopic degrees of freedom is a significant challenge, with considerable efforts focused on atomic structures generally [102–107] and GBs specifically [12–22]. In this work no new structural characterizations or machine learning methods are employed. Rather, we repeat methods employed previously by some of the authors to demonstrate the utility and wealth of data found in this dataset as it pertains to machine learning methods.

As noted in the supplemental material, atoms within ±15Å of the GB are available for analysis. In this work these atoms are analyzed using Smooth Overlap of Atomic Positions (SOAP) to characterize the local neighborhood of a selected atom [103]. SOAP is a descriptor that encodes regions of atomic geometries by using a local expansion of a Gaussian smeared atomic density with orthonormal spherical harmonic and radial basis functions. In this work we employ the SOAP implementation in the QUIPPY package [108–110], which outputs a vector of all the coefficients for the basis functions. The SOAP parameters used are as follows: $r_{\text{cut}} = 3.74$, $n_{\text{max}} = 12$, $l_{\text{max}} = 12$, $\sigma = 0.575$, *normalize* = true, and $Z = 13$. It is also worth noting that we do not include atoms within $r_{\text{cut}}$ of the ends of the simulation cells since their SOAP vectors do not match that of the bulk structure, despite their structure as bulk atoms.

Feature matrices for machine learning must always be the same size and, for traditional machine learning methods, each data point must be represented by a single vector. The SOAP analysis exports a SOAP matrix for each GB, comprised of SOAP vectors for each atom, that can be different in size since each GB can have different numbers of atoms in the atomic structure file. To obtain features that are the same size and to reduce the feature representation to single vector, the SOAP vectors in this matrix are averaged to provide one average SOAP vector that represents the GB atomic structure as a whole, and is referred to as the average SOAP representation (ASR) in [15].

While a matrix of these ASR vectors for each GB can be used as a feature matrix for machine learning, we choose to employ a kernel machine for the learning and must therefore construct a kernel from the ASR vectors. To be consistent with previous works [15], we elect to define our kernel matrix as $AA^T$ where $A$ is the ASR matrix. This results in a kernel matrix $K \in \mathbb{R}^{nxn}$ where $n$ is the number of features, 7304 in our case. The scikit-learn Python library is used to implement the support vector machine (SVM) algorithm [111]. The kernel from the ASR vectors essentially defines a similarity matrix between the GBs that is used by the SVM model to train and predict GB energy.

A second approach used for the machine learning in this work is referred to as the local environment representation (LER) [15]. The LER is a method to find a collection of "unique" local atomic environments by comparing the SOAP vectors of all GB atoms in the set and finding those that are unique as defined by a dissimilarity value $d$ less than a given magnitude $\epsilon$. The dissimilarity metric employed in this work is given by

$$d_{\vec{a},\vec{b}} = \left\| \vec{a} - \vec{b} \right\|_2 = \sqrt{\vec{a}\cdot\vec{a} + \vec{b}\cdot\vec{b} - 2\,\vec{a}\cdot\vec{b}} \qquad (1)$$

where $d_{\vec{a},\vec{b}}$ is the dissimilarity value when comparing the SOAP vectors $\vec{a}$ and $\vec{b}$ [17, 18].

Once a set of these unique local atomic environments are found for all atoms $r_{\text{cut}}$ away from the free surfaces in the set of 7304 GBs (50 247 007 atoms in total), the GB can then be represented by its fraction of each of these unique local atomic environments. The only parameter used to obtain the LER for the 7304 GBs in this work is $\epsilon = 0.05$. Using this parameter

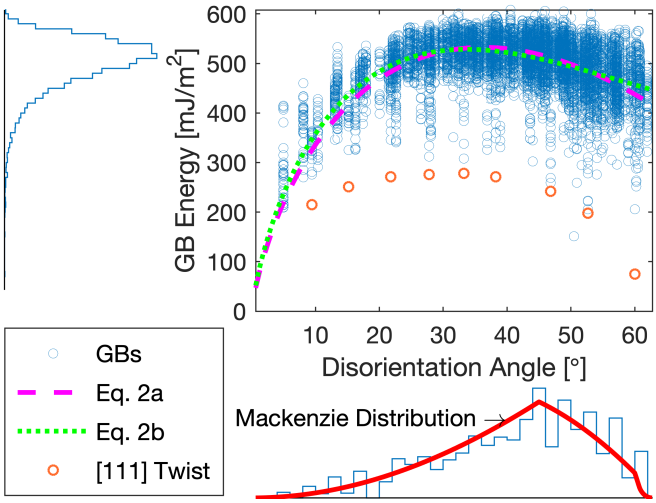


Figure 3: Scatter plot of energy as a function of disorientation angle for the 7304 GBs, accompanied by histograms of the two variables. Two possible Read-Shockley relationships are fit to the data and the low energy [1 1 1] disorientation axis twist GBs are also identified.

set, 1927 unique local atomic environments were found and the resulting LER matrix gives the fraction of each of these unique environments for each GB as rows in the matrix. Similar to ASR, a kernel matrix is constructed as $LL^T$ where L is the LER matrix. The LER kernel matrix is then used as the feature matrix and GB energy is predicted using SVM after the same manner as ASR.

## 3. Results & discussion

### 3.1. General dataset observations

The simplest manner to illustrate all the data in a single plot is to plot the minimum GB energy values as a function of a single variable, disorientation angle. This is shown in Fig. 3 for all 7304 GBs. The scatter plot of this data is accompanied by histograms of GB energy and the disorientation angle.

The range of GB energies compares well with those reported for an Olmsted set of Al GBs [54]. Supplemental Fig. S1 provides a comparison of GB energies for [111] Al GBs common to both the Olmsted dataset and the present work, calculated with the Ercolessi-Adams EAM potential [112] and Mishin et al. EAM potential [97], respectively. The GB energy values calculated using the Mishin potential are ~25 mJ/m$^2$ higher on average.

Figure 3 indicates a unimodal peak in the GB energy distribution, which has mean and standard deviation values of 497 ± 55 mJ/m$^2$. Note also the trend in the GB energy, which is similar to predictions made by Read-Shockley. In its most general sense, the Read-Shockley relationship can be defined as one of the following [113], where Eq. 2a is valid for low angles and Eq. 2b is valid for all angles.

$$\gamma(\theta)_{low} = \theta\left[E_c - E_s \ln(\theta)\right]/b \qquad (2a)$$

$$\gamma(\theta)_{all} = \sin(\theta)\left[E_c - E_s \ln(\sin(\theta))\right]/b \qquad (2b)$$

Table 2: Dislocation energy parameters and RMSE values for the two Read-Shockley relationships, along with comparison of expected dislocation energies as taken from the literature.

| Source | $E_c$ [eV/Å] | $E_s$ [eV/Å] | RMSE [mJ/m$^2$] |
|---|---|---|---|
| Eq. 2a | 0.082 | 0.150 | 45 |
| Eq. 2b | 0.070 | 0.170 | 45 |
| Literature* | 0.002-0.630 | 0.1893-0.2575 | - |

* Values for $E_c$ and $E_s$ are from [114] and [113], respectively.

The constants $E_c$ and $E_s$ represent the dislocation core energy and dislocation strain energy outside the core, respectively. These values are used as fitting parameters for the GB energy of the 7304 GBs in the dataset and their values are given in Table 2, along with the RMSE values from the fits. Fig. 3 has the fitted Eqs. 2a and 2b plotted on top of the individual data points. For reference, literature reports of dislocation core energy, $E_c$ are included [114], along with the range of $E_s$, based on screw and edge dislocations estimates, as given by Wolf [113] from which this form of the equations are taken. Note that the fit values are near expectations even though individual dislocations are not identifiable in many of the high angle GBs where dislocations would overlap.

It is noteworthy that an equation based on an idealized structure of dislocation arrays to mimic the structure and energy of GBs, developed over 70 years ago, captures the general behavior of a dataset that provides comprehensive coverage of the 5D space. While it is clear that these Read-Shockley fits do not describe all of the data points, the RMSE for both fits is 45 mJ/m$^2$. The utility of the Read-Shockley relationship is supported by other works as well, as indicated by Rohrer [34] and the fact that a modified form of Eq. 2b is used to interpolate GB energies between proximal boundaries in the Bulatov GB energy function [45], which has now be used to describe energies of four fcc and two bcc metals [45, 115, 116]. Finally, it is worth noting that the Read-Shockley fits peak between 30-40°and higher disorientation angle GBs have slightly lower GB energies. It should be pointed out however that at intermediate and high disorientation angles, dislocations overlap to the point that they cannot be readily distinguished. Thus, while the Read-Shockley model fits these higher disorientation angles, the interpretation of regularly spaced dislocations at this high angles, care should be exercised in the dislocation interpretation of the model.

The disorientation angle histogram in Fig. 3 is compared with the Mackenzie Distribution, which represents the distribution of disorientation angles for a sample with a random distribution of cubic crystal orientations [117, 118]. The similarity between the distribution of disorientation angles in the dataset and the Mackenzie distribution are indicative that the sampling of GBs in this dataset is similar to what would be observed if a sample with random cubic crystal orientations was analyzed. A histogram of just the 150 CSL disorientations is also provided in supplemental Fig. S2.

To make use of this data for GB engineering, one must understand how to control the texture, or populations of different crystal orientations. From Fig. 3 it is clear that a sample with random crystal orientations is unlikely to generate a GB distribution with many GBs that are low in energy. In fact, based on the distribution of energies in the histogram of Fig. 3, the fraction of low energy boundaries is bound to be negligible. Of course most GB engineering relies on the presence of twin boundaries and twin-related domains [2, 3, 119], which are typically limited to low stacking fault energy FCC materials. Aluminum does not belong in this group, but research suggests that alignment of {1 1 1}/{1 1 1} "interconnected" planes across GBs was frequently observed and has implications for GB engineering [120]. The {1 1 1} disorientation axis twist GBs, with {1 1 1}/{1 1 1} alignment, define the low energy bound as a function of disorientation angle in Fig. 3, with the exception of the $\Sigma$11 {1 1 3}/{1 1 3} symmetric tilt GB at a disorientation angle of 50.5°. In other words, this data supports the appearance of these {1 1 1}/{1 1 1} interconnected GBs, likely formed because of their low energy.

If we examine "low" energy GBs, defined by having an energy less than the highest {1 1 1} disorientation axis twist GBs, we find 27 GBs from 6 CSLs. Three near {1 1 1}/{1 1 1} GBs come from the $\Sigma$131$e$ (60.25°about [5 5 4]), which is vicinal to the $\Sigma$3, and therefore unsurprising to have low energy. 18 of these GBs come from the lowest disorientation angle GB studied in this work, the $\Sigma$265$a$ (4.98°about [1 0 0]), which is unsurprising considering low angle GBs are expected to have low energy. Another GB $\Sigma$201$a$ (8.09°about [1 1 0]), is also low angle, and therefore unsurprising to be included in this list. Finally, the $\Sigma$11 symmetric tilt GB, mentioned earlier, is well-known as a low energy GB. However, the remaining two GBs and CSLs are a bit surprising, these are a $\Sigma$63$c$ (54.03°about [4 3 1]) GB with BPs of (5 $\bar{1}$1 13)/(5 $\bar{1}$ 1$\bar{7}$) and a $\Sigma$69$d$ (50.92°about [5 5 1]) GB with BPs of (11 $\bar{1}$ 19)/(1 1$\bar{1}$ 1$\bar{9}$). To the authors' knowledge, these particular GBs and CSLs have not previously been identified as having low energy. Due to the coarse nature of the sampling of the 5D space, there may also be other low energy cusps that exist but which have not been discovered. In any case, the current dataset would suggest that beyond the three GBs mentioned above (from the $\Sigma$11, $\Sigma$63$c$, $\Sigma$69$d$ CSLs), an optimal texture would generate low angle GBs and {1 1 1}/{1 1 1} interconnected GBs.

The excess volume per unit area, referred to here as the excess volume, is also calculated for each GB. The excess volume, normalized by the lattice parameter, is plotted against GB energy in Fig. 4, along with a distribution of the excess volumes. The range of GB excess volumes compares well with values of excess volume in an Olmsted set of Al GBs [54]. Note the correlation of excess volume with GB energy (correlation coefficient of 0.8), which correlation has previously been shown to be linear [54]. The {1 1 1} disorientation axis twist GBs are again highlighted and have both low GB energy and low excess volume, as would be expected based on the correlation.

Also of interest in this dataset is a chance to examine the distribution of widths observed in the GB structures. In this work, we define the GB width by finding the GB atoms, identified by CNA values that are not FCC, and then measuring the dis-

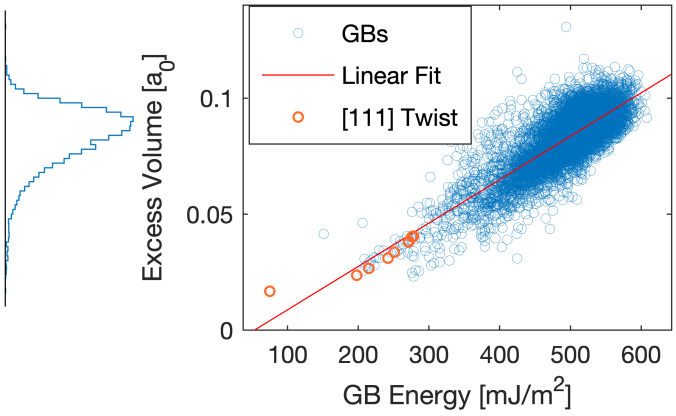


Figure 4: Scatter plot of excess volume, in units of the lattice parameter, as a function of GB energy for the 7304 GBs, accompanied by a histogram of the excess volume. The linear fit shown in the plot is given by $V_{ex} = 0.000187 \times E_{GB} - 0.00989$.

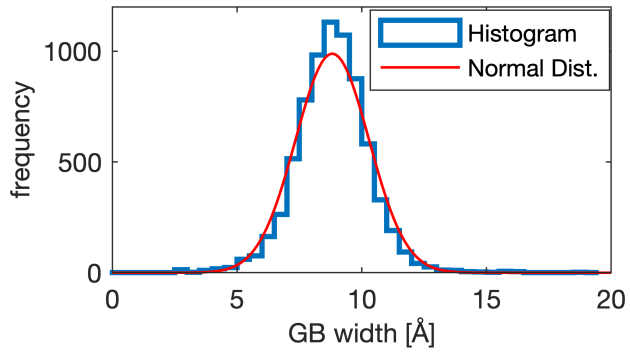


Figure 5: Histogram of GB widths compared with a normal distribution based on the mean and standard deviation of the data.

tance normal to the GB between the two most distant GB atoms on either side. Since it has been observed that defects, such as partial dislocations, can extend out from the GB some distance [81], this width really measures the extent of the disorder. It does not mean that all the atoms within the "GB width" are disordered. The distribution of GB widths is plotted in Fig. 5, where the mean and standard deviation of the distribution are $8.8 \pm 1.5$ Å. A Kolmogorov-Smirnov test indicates this distribution is normal with a significance value less than 0.001. These GB width values are in the range of estimates made by a variety of techniques in different materials [121–124].

### *3.2. GB energy trends in the 5D space*

One of the major challenges in examining GB structure-property relationships is the fact that the crystallographic character is 5-dimensional. This becomes challenging to graphically illustrate, which is why properties are often plotted as a function of only one variable, as in Fig. 3. Here we seek to demonstrate that GB energy varies smoothly through this 5D space, which has been demonstrated in previous cases, often over limited scopes or with limited data [29, 34, 45, 50–53, 58, 67, 71, 125, 126]. Fig. 6 plots the BP fundamental zones for several [1 1 1] disorientation axis CSLs with increasing disorientation angle. Note the similarity in GB energy trends and the general increase in GB energy with increasing disorientation angle. Note also the fact that the [1 1 1] twist GBs consistently have much lower energy values. In fact, these [1 1 1] twist GBs are the lowest points in Fig. 3.

Just as was done by Erickson and Homer [58], these BP fundamental zones can be stacked as a function of disorientation angle to make a quasi-3D volumetric plot of energy. The stacking occurs such that the twist and symmetric tilt GBs, which are at the vertices of the plots in Fig. 6, are consistently also at the vertices of the 3D volume. This is illustrated in Fig. 7 where all the GB datapoints are marked with an 'o' and isosurfaces of constant energy are included at various intervals to illustrate the gradient of energy between the sampled points. These isosurfaces illustrate how the energy increase as a function of disorientation angle is slower for the [1 1 1] twist GBs than the rest of the boundary planes. It can also be seen that just like the Read-Shockley equation predicts, at the highest disorientation angles, the GB energy decreases in magnitude.

While this work illustrates this variation along the [1 1 1] disorientation axis, Erickson and Homer illustrated the variation along the 1 0 0 disorientation axis in Ni [58]. These cases suggest that GB energy trends have similar trends for similar disorientation axes. This is supported by examinations of other disorientation axes where there were at least 3 different CSLs with the same disorientation axis. Supplemental Figs. S3-S6 show the isosurfaces of energy for these plots.

These evidences of smooth variation of GB energy in sections of the 5D space suggest that with the proper tools, one could examine trends of energy throughout the 5D space. Nearly all the 5D experimental datasets have examined GB energy throughout the 5D space and found repeated evidence that variation in BP orientation is associated with greater differences in GB energy than variation in disorientation [34]. This is supported by the present data as can be seen by the vertical lines of data points in Fig. 3 and the variation of energy in Figs. 6 and 7. Additionally, Bulatov et al. in their GB energy function, note that the cusps and grooves follow the natural symmetries of the underlying lattice [45]. This can be seen in part by the fact that the low and high energy values in Figs. 6 and 7 occur at the natural symmetry points or vertices of the BP fundamental zone. Some of these features can be seen in [29, 58] and in the supplemental Figs. S3-S6.

Recent developments in this area show great potential to understand this space further, with new distance metrics, interpolation schemes, expansions, visualizations, and other methods, to examine property GB character and property variations throughout the 5D space. For example, Francis et al. developed the octonion to compare and interpolate between GBs [127]. This was followed by a number of tools to examine the GB manifold [128]. Baird et al. recently demonstrated an efficient interpolation technique in the 5D space using a Voronoi fundamental zone framework [85] as well as methods for quantitative cartography of the GB energy landscape [126]. Hu et al. used a genetic algorithm and deep neural network to construct 5DOF property diagrams [129]. Mason and Patala recently introduced a new construction using orthornormal basis functions

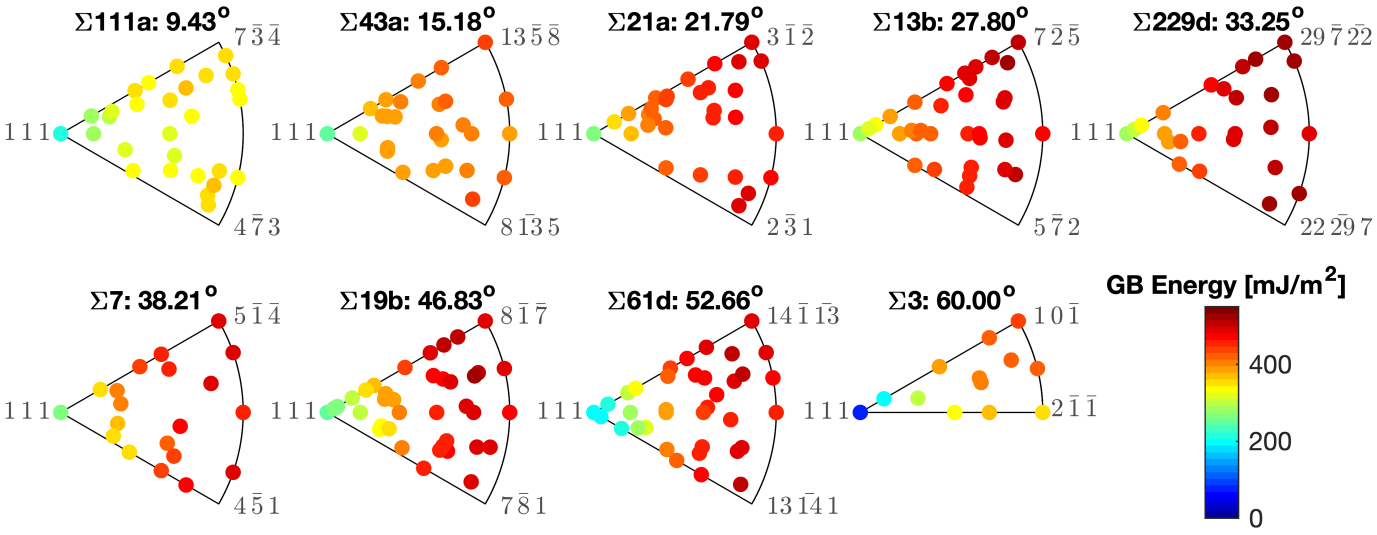


Figure 6: Plots of GB energy in the BP fundamental zones for 9 different [1 1 1] disorientation axis CSLs. Note the similar trend of increasing energy from the [1 1 1] twist GB toward the [1 1 1] tilt GBs along the arc.

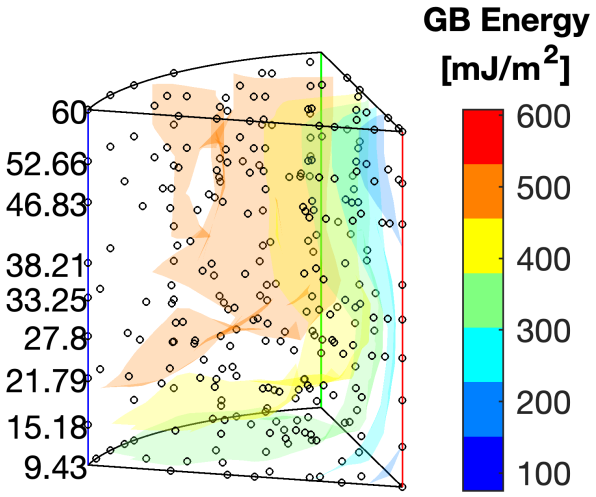


Figure 7: Disorientation angle-BP volumetric plot with isosurfaces of different GB energy values plotted to show the variation within the space of [1 1 1] disorientation axis GBs. This plot is essentially a vertical stacking of the images in Fig. 6.

that account for crystallographic point group symmetries, grain exchange symmetry, and the null boundary singularity [130]. This sampling of recent works demonstrates that the community is poised to gain insight into the 5D space that was not possible before. It is hoped that the comprehensive 5D sampling provided by the present work will be a useful tool in gaining this new insight.

### *3.3. Metastability*

As mentioned earlier, while producing this dataset, we stored the GB energy and structure for all structures that were minimized, not just the minimum energy case. In all, this amounted to 43 009 236 structures to obtain the minimum energy structures for the 7304 GBs. While these non-minimum energy structures are not included with the published dataset at this time, we present here a few insights into the energies of these additional structures. Fig. 8a provides a plot of all 43 009 236 energies as a function of disorientation angle, along with a histogram of the energies. It can be seen that the distribution of energies in the histogram is not significantly altered from Fig. 3. The mean and standard deviation of the distribution are 546±55 mJ/m$^2$. This is the same standard deviation but the mean is about 50 mJ/m$^2$ higher.

Interestingly, the [1 1 1] disorientation axis twist GBs that have the lowest energy are also the most likely to have the highest possible energy from one of their structures; this can be seen from the fact that the vertical array of data points for these low energy GBs also extend well above the majority of boundaries in the energy distribution. We choose to analyze the twist and two symmetric tilt GBs from the [1 1 1] disorientation axis GBs, that are shown in Figs. 6 and 7. In fact these correspond to the GBs whose normals are labeled at two of the three vertices as well as the mid-point of the arc. Swarm plot distributions for these [1 1 1] symmetric tilt and twist GBs are shown in Fig. 8b. Here it is obvious that the twist GBs have both the lowest and highest GB energies. In contrast, the two symmetric tilt GB energies cover a much smaller range. For the whole dataset, the average and standard deviation of the range of energies covered by the structures for each of the 7304 GBs is 152 ± 67 mJ/m$^2$.

It is worth noting that work by Foley and Tucker examining damage tolerance of GBs follows a similar trend [131]. In their work, the lowest energy GBs would absorb significantly more defects and increase the free volume in the GB before achieving an equilibrium state at a much higher GB energy. In contrast, the GBs that started with a higher energy absorbed a smaller number of defects and increased the free volume in the GB by a smaller amount before achieving an equilibrium state at a modest increase to the GB energy. In short, the high energy GBs acted as more efficient sinks for defects, whereas the low energy GBs were not efficient sinks and accommodating defects

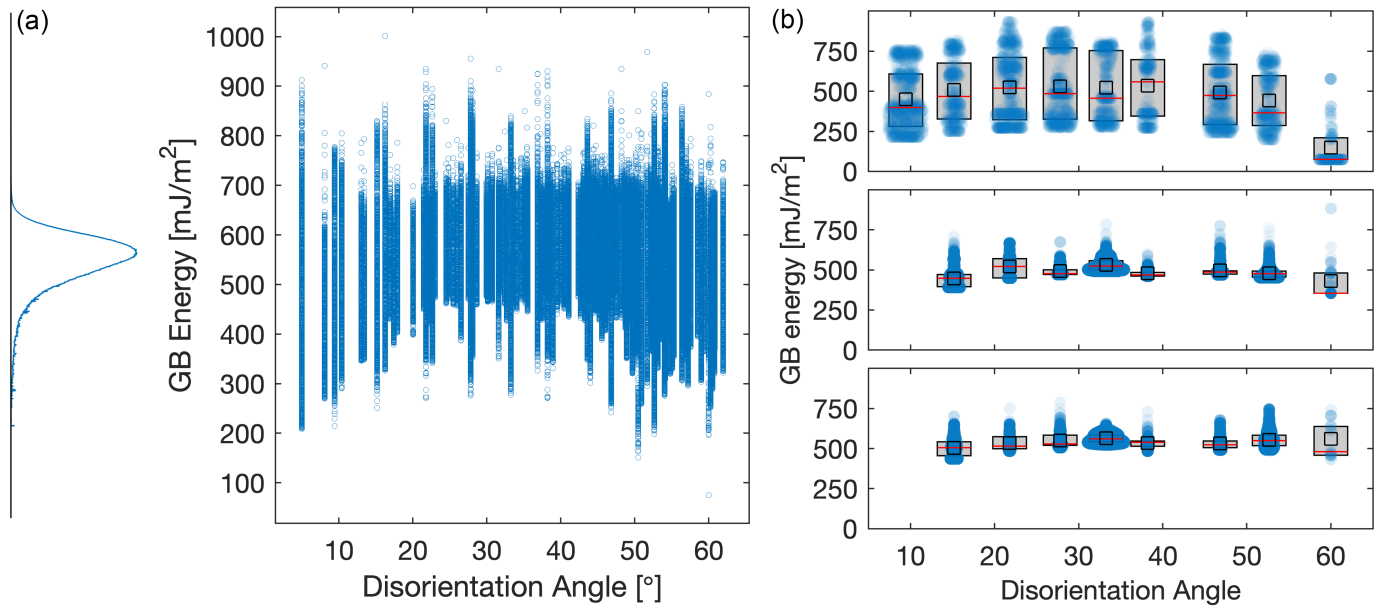


Figure 8: (a) Scatter plot of GB energy as a function of disorientation angle for all 43 009 236 minimized structures, accompanied by histograms of the GB energy. (b) Swarm plots of the GB energy for the [1 1 1] twist, and the two [1 1 1] symmetric tilt GBs from top to bottom, respectively. The swarm plots are overlaid with a box marking the first and third quartiles of the distribution, the red line marks the median, and the black square marks the mean.

resulted in large increases to the GB energy. If these [1 1 1] disorientation axis twist GBs were to be advantageous because they have lower energy, in an environment where defect accumulation might be required (e.g., irradiation), the GB would quickly turn into a high energy GB, with possible energies in excess of that of the defect free high energy GBs, and maybe even in excess of their accumulated defect state.

It must also be mentioned at this point that we are unsure which of all these structures actually represent unique metastable states. For example, in Fig. 8b it can be seen that many of the swarm plots have flat distributions at the bottom. This is representative of the fact that many of the starting configurations, although unique as an input file, found the same minimum energy configuration.

Previous methods to find metastable states [25] used a finer sampling of 2D grain translations but did not consider other construction variables, such as boundary placement, which allowed them to determine which states were unique. In order to provide comprehensive coverage of the 5D space, we were not able to sample in such a manner. Methods such as USPEX [132, 133] are able to determine unique metastable states as the GB structure evolves [14]. Zhu et al. used machine learning to determine the common structural features of GBs [14] and such a method may prove fruitful in the future to determine which of all these structures represent truly unique metastable structures.

If one can indeed determine which of all these structures represent unique metastable states, a number of potential research directions become available. First off, one can perform statistical mechanical calculations to determine expected values of different properties, as done by Han et al. [25]. It is also possible that these metastable states represent the states through which a GB would pass as it migrates or deforms, as done by Alexander and Schuh [134, 135]. Such an approach would allow prediction of dynamic properties of a GB. Winter et al. recently devised a theory for nucleation of different GB phases [136]. Finally, with knowledge about the energy of different metastable states, and even better if they could be connected into a potential energy landscape, one could potentially predict thermodynamic properties, such as chemical potential, which is defined by the change in energy with respect to changes in the number of atoms. Thus, the metastable GB structures presented here have significant potential in advancing our understanding of GBs.

### *3.4. Machine learning*

Here we turn our attention to the atomic structure of the GBs, which represent the microscopic degrees of freedom of a GB. As described in the methods (2.3), we characterize all the atomic structures of the 7304 GBs using both ASR and LER. Also mentioned in the methods is that, in this work, we make no effort to innovate in characterization or machine learning approaches. Rather, the focus here is to illustrate the richness of this dataset as it pertains to potential uses in machine learning.

As noted in the methods, following ASR characterization, a kernel matrix is defined that represents the similarity between all the GBs. This kernel matrix is illustrated in supplemental Fig. S7 where it can be seen that the the range of values in the kernel is quite small. Clearly some GBs are more like and unlike other GBs, but the interesting thing to note is that the feature matrix does not appear to provide significant discrimination.

With the feature matrix defined by the kernel matrix, we run support vector regression with various splits of the data into training and validation sets. These splits are defined in Table 3. In the first split, a 5-fold cross-validation is carried out with a random split of the GBs. Since a random split could allow

Table 3: Comparison of machine learning accuracy for ASR and LER as measured by RMSE and $R^2$ for a variety of training and validation splits of the dataset. For the answers with 5-Fold cross-validation (CV), the RSME and $R^2$ are given as the mean ± the standard deviation. The low and high disorientation angle split is defined at a disorientation angle of 15°, with 250 and 7054 GBs in the groups, respectively.

| Characterization and Dataset Split | RMSE [mJ/m$^2$] | | $R^2$ | |
|---|---|---|---|---|
| ASR | Train | Validation | Train | Validation |
| 5-Fold CV Random Split | 12.85±0.09 | 12.88±0.33 | .945±0.001 | .945±0.004 |
| 5-fold CV Disjoint Disorientation Angle Split | 10.45±1.10 | 13.17±1.15 | .94±.03 | .94±.01 |
| Disorientation Angle Train Low-Validation High | 24.15 | 24.09 | .87 | .77 |
| Disorientation Angle Train High-Validation Low | 11.28 | 13.55 | .94 | .94 |
| Shuffled GB energies Train 2/3-Validation 1/3 | 56.58 | 54.64 | -.037 | -.036 |
| LER | Train | Validation | Train | Validation |
| 5-Fold CV Random Split | 23.90±0.11 | 14.05±0.42 | .95±0.001 | .93±0.003 |
| 5-fold CV Disjoint Disorientation Angle Split | 5.32±.27 | 16.81±.32 | .96±.006 | .91±.007 |
| Disorientation Angle Train Low-Validation High | 15.71 | 33.91 | .94 | .54 |
| Disorientation Angle Train High-Validation Low | 14.92 | 18.70 | .91 | .92 |
| Shuffled GB energies Train 2/3-Validation 1/3 | 54.60 | 55.64 | .032 | -.073 |

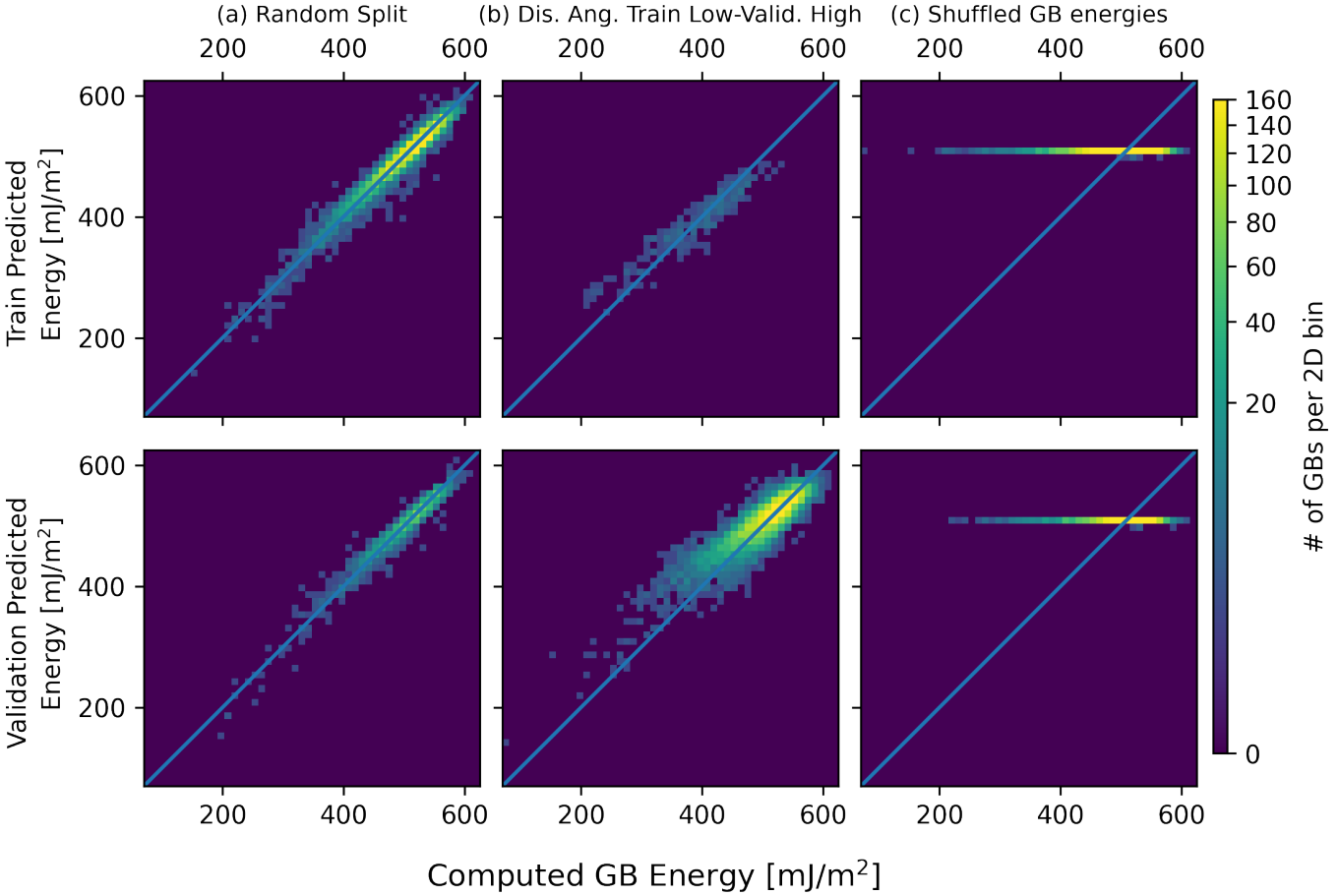


Figure 9: Parity plots comparing predicted vs. computed GB energy for both the training and validation sets trained on ASR. The various splits match some of those described in Table 3, and are identified as (a) one of the 5-Fold cross-validation random splits, (b) the disorientation angle train low-validation high split, and (c) the shuffled GB energies train 2/3-validation 1/3 split.

training of the model at nearby points, we perform a 5-fold split that is disjoint over the disorientation angles, ensuring that each split has approximately the same number of boundaries. We also split on disorientation angle into low and high angle GBs, with the split at 15°. We perform two tests, training on the low angle GBs and validation on the high angle GBs and vice versa. Finally, for a control, we perform a random shuffling of the GB energy values and train and validate the model.

The results of all these tests are shown in Table 3 as measured by the RMSE and $R^2$ values. This is accompanied by Fig. 9 which has 2D visualizations comparing the computed and training predictions of GB energy and the computed and validation predictions of GB energy for 3 of the splits listed in Table 3. In Fig. 9, a perfect prediction either in the training or validation would fall along the parity line. The first column has one of the 5 folds in the random split for the cross validation, where the predictions are good as evidenced by both the RMSE and $R^2$ values. The disjoint disorientation angle split also has low RMSE and high $R^2$ values.

The train low-validate high disorientation angle split is illustrated in the second column where only 250 low angle GBs are used to train the model and predictions are made on the 7054 validation GBs. The validation RMSE value is about double the previous examples and the $R^2$ value is lower, but still indicative that the majority of the variance in the predicted energy is accounted for in the model.

Finally, the last column in Fig. 9 has a random shuffling of the GB energy values such that there should be no correlation to be found in the machine learning. In the training and validation, the $R^2$ value near zero is indicative that the machine learning finds no correlation, as expected. This is clearly evident in the figure, and the predictions are just over the mean value of the GB energy shown in Fig. 3 of 497 mJ/m$^2$ and the RMSE value is about equal to the standard deviation in Fig. 3 of 55 mJ/m$^2$ for both the training and validation. Thus, the random shuffle is only able to predict the mean, but the large concentration of the GB energy about the mean value means that the RMSE values are low. This also means that we must take care in interpreting the RMSE values.

The same machine learning process is repeated with the LER kernel matrix and, similar to previous works [15], the LER learning is not as good as the ASR learning, as evidenced by the RMSE and $R^2$ values in Table 3. While the initial goal of LER was to get away from the averaging and information loss that occurs in ASR, it is notable that the predictions from ASR are consistently better. The reason for this may be that while LER attempts to get away from a single "average" description in favor of representing GB structure from a distribution of unique local atomic environments, it discards all information about those unique atomic environments calculated with SOAP. This is then reduced further in the formation of the kernel, thus losing additional information.

To reiterate, the goal of showing this learning is to demonstrate that this dataset will be useful for machine learning approaches to find structure property relationships of GBs that can focus on the microscopic (atomic) degrees of freedom. So, while these two approaches provide noteworthy learning and validation, the challenge remains to find feature representations that are consistent in size and also encode relevant information about the GB structure for learning on a variety of properties and phenomena. We are optimistic that this dataset will provide sufficient data for individuals to devise and test clever characterization and machine learning approaches to advance GB structure-property relationships that span the full 5D crystallographic (macroscopic) character while taking into account the atomic (microscopic) degrees of freedom that define the structure and properties of GBs.

## 4. Conclusions

This work examines a computed dataset of aluminum GBs that provide comprehensive coverage of the 5D crystallographic character space. The dataset is comprised of 7304 GBs that cover a range of BPs in 150 different CSLs that span disorientation space. For each GB, a number of possible structural configurations is considered and minimized. This results in more than 43 million possible structures that could be metastable states corresponding to the 7304 GBs. The following observations are made concerning this dataset, with a particular focus on the minimum energy GB structures and properties.

- The distribution of disorientation angles in the dataset match a Mackenzie distribution, indicating equivalence to GBs that would be present in a sample with random cubic crystal orientations.
- The distribution of GB energies in the dataset is unimodal with a mean and standard deviation of 497 ± 55 mJ/m$^2$.
- The GB energies as a function of disorientation angle follow the general trend of the Read-Shockley relationship with a peak between 30-40°, with a subsequent decrease in energy for the highest disorientation angles.
- The {1 1 1}/{1 1 1} interconnected GBs have the lowest GB energies for all disorientation angles, with the exception of the $\Sigma$11 {1 1 3}/{1 1 3} symmetric tilt GB. This is consistent with experimental observations where these {1 1 1}/{1 1 1} interconnected GBs are expected to have low energy.
- The GB energy and excess volume per unit area exhibit a linear correlation.
- The GB width, as measured by the extent of disorder normal to the GB, has a normal distribution with a mean and standard deviation of 8.8 ± 1.5 Å.
- The GB energy, as examined in subspaces of the 5D space, show smooth variation and similar behaviors for similar disorientation axes.
- Examination of all the GB energy values obtained during the sampling of the 43 009 236 structures is also Read-Shockley-like, and the distribution of values is unimodal with a mean and standard deviation of 546 ± 55 mJ/m$^2$.
- Examination of all the GB energy values obtained during the sampling of the [1 1 1] disorientation axis twist GBs, which are the same {1 1 1}/{1 1 1} interconnected GBs, have the largest spread in GB energies for the non-minimum energy structures, ranging from the lowest possible energies at a given disorientation angle to the highest possible GB energies. In contrast the [1 1 1] disorientation axis symmetric tilt GBs, which have higher GB energy,

have a much smaller range. This is consistent with observations of damage accumulation where low energy GBs are unable to accommodate defects without a significant increase in GB energy.

- Machine learning of GB atomic structures to predict GB energy using previously published methods, referred to as average SOAP representation (ASR) and local environment representation (LER), perform well on training and validation sets as measured by RMSE and $R^2$ values. This is true even when training on small datasets, such as training on 250 low angle GBs and predicting on all 7054 high angle GBs.

In some ways, this dataset was unsurprising and reinforced many known facts about GBs. For example, the fact that the {1 1 1}/{1 1 1} interconnected GBs have the lowest energies is unsurprising. Also, only a few new GBs were found with comparably low energies. However, the dataset provides insight about the variation of properties and structure across the 5D space of crystallographic character, which is information that is critical to microstructure sensitive design and GB engineering.

The authors are optimistic that this dataset will be useful for learning GB structure-property relationships on both the 5D crystallographic (macroscopic) character as well as the atomic (microscopic) structure of GBs. As an example, a large database of handwritten numbers, known as MNIST, has long provided fertile ground for training of image processing methods and other machine learning models and serves as a basis for benchmarking different algorithms [137–139]. With the explosion of methods to examine GBs at both the macroscopic and microscopic scales, this dataset may prove useful in a similar way.

## Dataset

The dataset for this work is published on Mendeley Data [87], with an embargo date of 12/31/2022. Individuals interested in obtaining the dataset prior to this date may contact the authors. The dataset includes a CSV file with details about the crystallographic character, calculated properties, and other information about the 7304 GBs in the dataset. The CSV file is accompanied by a zipped directory containing the atomic structures of all the minimum energy configurations for the 7304 GBs in this work.

## Acknowledgments

This work was supported by the U.S. National Science Foundation (NSF) under Award #DMR-1817321.

## References

[1] G. Palumbo, E. Lehockey, P. Lin, Applications for grain boundary engineered materials, JOM-Journal Of The Minerals Metals & Materials Society 50 (2) (1998) 40–43.

[2] T. Watanabe, S. Tsurekawa, X. Zhao, L. Zuo, The coming of grain boundary engineering in the 21st century, in: A. Haldar, S. Suwas, D. Bhattacharjee (Eds.), Microstructure and Texture in Steels, Springer London, London, 2009, pp. 43–82.

[3] V. Randle, Grain boundary engineering: an overview after 25 years, Materials Science And Technology 26 (3) (2010) 253–261. doi:10.1179/026708309X12601952777747.

[4] P. Lin, G. Palumbo, U. Erb, K. T. Aust, Influence of grain boundary character distribution on sensitization and intergranular corrosion of alloy 600, Scripta Metallurgica et Materialia 33 (9) (1995) 1387–1392. doi:10.1016/0956-716X(95)00420-Z.

[5] E. M. Lehockey, D. Limoges, G. Palumbo, J. Sklarchuk, K. Tomantschger, A. Vincze, On improving the corrosion and growth resistance of positive Pb-acid battery grids by grain boundary engineering, Journal of power sources 78 (1) (1999) 79 – 83. doi:10.1016/S0378-7753(99)00015-4.

[6] V. Thaveeprungsriporn, G. S. Was, The role of coincidence-site-lattice boundaries in creep of Ni-16Cr-9Fe at 360 °C, Metallurgical and Materials Transactions A 28 (10) (1997) 2101–2112. doi:10.1007/s11661-997-0167-6.

[7] S. Bechtle, M. Kumar, B. Somerday, M. Launey, R. Ritchie, Grain-boundary engineering markedly reduces susceptibility to intergranular hydrogen embrittlement in metallic materials, Acta Materialia 57 (14) (2009) 4148–4157. doi:10.1016/j.actamat.2009.05.012.

[8] R. W. Balluffi, P. D. Bristowe, On the structural unit/grain boundary dislocation model for grain boundary structure, Surface Science 144 (1) (1984) 28–43. doi:10.1016/0039-6028(84)90701-5.

[9] A. P. Sutton, On the structural unit model of grain boundary structure, Philosophical Magazine Letters 59 (2) (1989) 53–59.

[10] J. Han, V. Vitek, D. J. Srolovitz, The grain-boundary structural unit model redux, Acta Materialia 133 (2017) 186–199. doi:10.1016/j.actamat.2017.05.002.

[11] A. D. Banadaki, S. Patala, A three-dimensional polyhedral unit model for grain boundary structure in fcc metals, npj Computational Materials 3 (1) (2017) 13. doi:10.1038/s41524-017-0016-0.

[12] T. Tamura, M. Karasuyama, R. Kobayashi, R. Arakawa, Y. Shiihara, I. Takeuchi, Fast and scalable prediction of local energy at grain boundaries: machine-learning based modeling of first-principles calculations, Modelling And Simulation In Materials Science And Engineering 25 (7) (2017) 075003. doi:10.1088/1361-651X/aa8276.

[13] J. A. Gomberg, A. J. Medford, S. R. Kalidindi, Extracting knowledge from molecular mechanics simulations of grain boundaries using machine learning, Acta Materialia 133 (2017) 100–108. doi:10.1016/j.actamat.2017.05.009.

[14] Q. Zhu, A. Samanta, B. Li, R. E. Rudd, T. Frolov, Predicting phase behavior of grain boundaries with evolutionary search and machine learning, Nat. Commun. 9 (1). doi:10.1038/s41467-018-02937-2.

[15] C. W. Rosenbrock, E. R. Homer, G. Csányi, G. L. W. Hart, Discovering the building blocks of atomic systems using machine learning: application to grain boundaries, npj Computational Materials 3 (1) (2017) 29. doi:10.1038/s41524-017-0027-x.

[16] T. A. Sharp, S. L. Thomas, E. D. Cubuk, S. S. Schoenholz, D. J. Srolovitz, A. J. Liu, Machine learning determination of atomic dynamics at grain boundaries, Proc. Natl. Acad. Sci. 115 (43) (2018) 10943–10947. doi:10.1073/pnas.1807176115.

[17] C. W. Rosenbrock, J. L. Priedeman, G. L. W. Hart, E. R. Homer, Structural characterization of grain boundaries and machine learning of grain boundary energy and mobility (2018). arXiv:1808.05292.

[18] J. L. Priedeman, C. W. Rosenbrock, O. K. Johnson, E. R. Homer, Quantifying and connecting atomic and crystallographic grain boundary structure using local environment representation and dimensionality reduction techniques, Acta Materialia 161 (2018) 431–443. doi:10.1016/j.actamat.2018.09.011.

[19] B. D. Snow, D. D. Doty, O. K. Johnson, A Simple Approach to Atomic Structure Characterization for Machine Learning of Grain Boundary Structure-Property Models, Frontiers in Materials 6 (2019) 054104. doi:10.3389/fmats.2019.00120.

[20] S. Patala, Understanding grain boundaries – the role of crystallography, structural descriptors and machine learning, Computational Materials Science 162 (2019) 281–294. doi:10.1016/j.commatsci.2019.02.047.

[21] E. R. Homer, D. M. Hensley, C. W. Rosenbrock, A. H. Nguyen,

G. L. W. Hart, Machine-Learning Informed Representations for Grain Boundary Structures, Frontiers in Materials 6 (2019) 168. doi:10.3389/fmats.2019.00168.

[22] David Montes de Oca Zapiain, Matthew Guziewski, Shawn P. Coleman, Rémi Dingreville, Characterizing the tensile strength of metastable grain boundaries in silicon carbide using machine learning, The Journal of Physical Chemistry C 124 (45) (2020) 24809–24821. doi:10.1021/acs.jpcc.0c07590.

[23] E. W. Hart, Grain Boundary Phase Transformations, in: H. Hu (Ed.), The Nature and Behavior of Grain Boundaries, Springer, New York, 1972, pp. 155–170. doi:10.1007/978-1-4757-0181-4_6.

[24] T. Frolov, D. L. Olmsted, M. Asta, Y. Mishin, Structural phase transformations in metallic grain boundaries, Nature Communications 4 (2013) 1899. doi:10.1038/ncomms2919.

[25] J. Han, V. Vitek, D. J. Srolovitz, Grain-boundary metastability and its statistical properties, Acta Materialia 104 (2016) 259–273. doi:10.1016/j.actamat.2015.11.035.

[26] J. Hickman, Y. Mishin, Extra variable in grain boundary description, Physical Review Materials 1 (1) (2017) 010601. doi:10.1103/PhysRevMaterials.1.010601.

[27] T. Meiners, T. Frolov, R. E. Rudd, G. Dehm, C. H. Liebscher, Observations of grain-boundary phase transformations in an elemental metal, Nature 579 (7799) (2020) 375 – 378. doi:10.1038/s41586-020-2082-6.

[28] S. Patala, C. A. Schuh, Symmetries in the representation of grain boundary-plane distributions, Philosophical Magazine 93 (5) (2013) 524–573. doi:10.1080/14786435.2012.722700.

[29] E. R. Homer, S. Patala, J. L. Priedeman, Grain Boundary Plane Orientation Fundamental Zones and Structure-Property Relationships , Scientific Reports 5 (2015) 15476. doi:10.1038/srep15476.

[30] H. Gleiter, On the structure of grain boundaries in metals, Materials Science And Engineering 52 (2) (1982) 91–131. doi:10.1016/0025-5416(82)90040-4.

[31] D. Wolf, S. Yip (Eds.), Materials Interfaces, Atomic-level Structure and Properties, Chapman & Hall, London, 1992.

[32] A. Sutton, R. Balluffi, Interfaces in Crystalline Materials, Oxford University Press, Oxford, 1995.

[33] G. Gottstein, L. S. Shvindlerman, Grain Boundary Migration in Metals, CRC Press, Boca Raton, 2010.

[34] G. S. Rohrer, Grain boundary energy anisotropy: a review, Journal Of Materials Science 46 (18) (2011) 5881–5895. doi:10.1007/s10853-011-5677-3.

[35] W. T. Read, W. Shockley, Dislocation Models of Crystal Grain Boundaries, Physical Review 78 (3) (1950) 275–289. doi:10.1103/PhysRev.78.275.

[36] D. Wolf, Structure and energy of general grain boundaries in bcc metals, Journal Of Applied Physics 69 (1) (1991) 185–196.

[37] D. Wolf, A broken-bond model for grain boundaries in face-centered cubic metals, Journal Of Applied Physics 68 (7) (1990) 3221–3236.

[38] J. W. Cahn, Y. Mishin, A. Suzuki, Duality of dislocation content of grain boundaries, Philosophical Magazine 86 (25-26) (2006) 3965–3980. doi:10.1080/14786430500536909.

[39] J. B. Yang, Y. Nagai, M. Hasegawa, Use of the Frank–Bilby equation for calculating misfit dislocation arrays in interfaces, Scripta Materialia 62 (7) (2010) 458–461. doi:10.1016/j.scriptamat.2009.12.011.

[40] A. Sangghaleh, M. J. Demkowicz, AIDA: A tool for exhaustive enumeration of solutions to the quantized Frank-Bilby equation, Computational Materials Science 145 (2018) 35–47. doi:10.1016/j.commatsci.2017.12.025.

[41] D. W. Hoffman, J. W. Cahn, Vector Thermodynamics for Anisotropic Surfaces .1. Fundamentals and Application to Plane Surface Junctions, Surface Science 31 (1) (1972) 368–388.

[42] J. W. Cahn, D. W. Hoffman, Vector Thermodynamics for Anisotropic Surfaces .2. Curved and Faceted Surfaces, Acta Metallurgica Et Materialia 22 (10) (1974) 1205–1214.

[43] A. A. Wheeler, Cahn-Hoffman xi-vector and its relation to diffuse interface models of phase transitions, Journal of Statistical Physics 95 (5-6) (1999) 1245–1280.

[44] R. W. Balluffi, S. Allen, W. C. Carter, Kinetics of Materials, John Wiley & Sons, Hoboken, 2005.

[45] V. V. Bulatov, B. W. Reed, M. Kumar, Grain boundary energy function for fcc metals, Acta Materialia 65 (2014) 161–175. doi:10.1016/j.actamat.2013.10.057.

[46] H. J. Frost, M. F. Ashby, F. Spaepen, A Catalogue of [100], [110], and [111] Symmetric Tilt Boundaries in Face-Centered Cubic Hard Sphere Crystals, Harvard Division of Applied Sciences (1982) 1–216.

[47] J. E. Spowart, Automated serial sectioning for 3-D analysis of microstructures, Scripta Materialia 55 (1) (2006) 5–10. doi:10.1016/j.scriptamat.2006.01.019.

[48] S. J. Dillon, G. S. Rohrer, Characterization of the Grain-Boundary Character and Energy Distributions of Yttria Using Automated Serial Sectioning and EBSD in the FIB, Journal Of The American Ceramic Society 92 (7) (2009) 1580–1585. doi:10.1111/j.1551-2916.2009.03064.x.

[49] M. P. Echlin, T. L. Burnett, A. T. Polonsky, T. M. Pollock, P. J. Withers, Serial sectioning in the SEM for three dimensional materials science, Current Opinion in Solid State and Materials Science 24 (2) (2020) 100817. doi:10.1016/j.cossms.2020.100817.

[50] D. Wolf, Structure-energy correlation for grain boundaries in F.C.C. metals—I. Boundaries on the (111) and (100) planes, Acta Metallurgica Et Materialia 37 (7) (1989) 1983–1993. doi:10.1016/0001-6160(89)90082-5.

[51] D. Wolf, Structure-energy correlation for grain boundaries in f.c.c. metals—II. Boundaries on the (110) and (113) planes, Acta Metallurgica Et Materialia 37 (10) (1989) 2823–2833.

[52] D. Wolf, Structure-Energy Correlation for Grain Boundaries in FCC metals—III. Symmetrical Tilt Boundaries, Acta metallurgica et materialia 38 (5) (1990) 781–790. doi:10.1016/0956-7151(90)90030-K.

[53] D. Wolf, Structure-Energy Correlation for Grain Boundaries in FCC Metals—IV. Asymmetrical Twist (General) Boundaries, Acta metallurgica et materialia 38 (5) (1990) 791–798.

[54] D. L. Olmsted, S. M. Foiles, E. A. Holm, Survey of computed grain boundary properties in face-centered cubic metals: I. Grain boundary energy, Acta Materialia 57 (13) (2009) 3694–3703. doi:10.1016/j.actamat.2009.04.007.

[55] S. Ratanaphan, D. L. Olmsted, V. V. Bulatov, E. A. Holm, A. D. Rollett, G. S. Rohrer, Grain boundary energies in body-centered cubic metals, Acta Materialia 88 (C) (2015) 346–354. doi:10.1016/j.actamat.2015.01.069.

[56] M. A. Tschopp, S. P. Coleman, D. L. McDowell, Symmetric and asymmetric tilt grain boundary structure and energy in Cu and Al (and transferability to other fcc metals), Integr. Mater. Manuf. Innov. 4 (1) (2015) 176–189. doi:10.1186/s40192-015-0040-1.

[57] H. C. Erickson, J. L. Priedeman, E. R. Homer, [100] disorientation axis grain boundary energies and structures, Mendeley Data (2020). doi:10.17632/bz46sgxm8g.

[58] H. C. Erickson, E. R. Homer, Insights into grain boundary energy structure-property relationships by examining computed [100] disorientation axis grain boundaries in nickel, Scripta Materialia 185 (2020) 165–169. doi:10.1016/j.scriptamat.2020.03.062.

[59] M. Guziewski, A. D. Banadaki, S. Patala, S. P. Coleman, Application of Monte Carlo techniques to grain boundary structure optimization in silicon and silicon-carbide, Computational Materials Science 182 (2020) 109771. doi:10.1016/j.commatsci.2020.109771.

[60] Y. Zhang, E. D. Hansen, T. Harbison, S. Masengale, J. French, L. Aagesen, A molecular dynamics survey of grain boundary energy in uranium dioxide and cerium dioxide, Journal of the American Ceramic Society-doi:10.1111/jace.18340.

[61] H.-K. Kim, W.-S. Ko, H.-J. Lee, S. G. Kim, B.-J. Lee, An identification scheme of grain boundaries and construction of a grain boundary energy database, Scripta Materialia 64 (12) (2011) 1152–1155. doi:10.1016/j.scriptamat.2011.03.020.

[62] H. Zheng, X.-G. Li, R. Tran, C. Chen, M. Horton, D. Winston, K. A. Persson, S. P. Ong, Grain boundary properties of elemental metals, Acta Materialia 186 (2020) 40–49. doi:10.1016/j.actamat.2019.12.030.

[63] B. L. Adams, S. Ta'asan, D. Kinderlehrer, I. Livshits, D. Mason, C. Wu, W. Mullins, G. S. Rohrer, A. D. Rollett, D. Saylor, Extracting grain boundary and surface energy from measurement of triple junction geometry, Interface Science 7 (3-4) (1999) 321–338.

[64] A. Morawiec, Method to calculate the grain boundary energy distribution over the space of macroscopic boundary parameters from the geometry of triple junctions, Acta Materialia 48 (13) (2000) 3525–3532.

[65] D. M. Saylor, A. Morawiec, G. S. Rohrer, The relative free energies of grain boundaries in magnesia as a function of five macroscopic param-

eters, Acta Materialia 51 (13) (2003) 3675–3686. doi:10.1016/S1359-6454(03)00182-4.

[66] J. Li, S. Dillon, G. Rohrer, Relative grain boundary area and energy distributions in nickel, Acta Materialia 57 (14) (2009) 4304–4311. doi:10.1016/j.actamat.2009.06.004.

[67] D. M. Saylor, B. S. El Dasher, A. D. Rollett, G. S. Rohrer, Distribution of grain boundaries in aluminum as a function of five macroscopic parameters, Acta Materialia 52 (12) (2004) 3649–3655. doi:10.1016/j.actamat.2004.04.018.

[68] E. A. Holm, G. S. Rohrer, S. M. Foiles, A. D. Rollett, H. M. Miller, D. L. Olmsted, Validating computed grain boundary energies in fcc metals using the grain boundary character distribution, Acta Materialia 59 (13) (2011) 5250–5256. doi:10.1016/j.actamat.2011.05.001.

[69] Y. Amouyal, E. Rabkin, Y. Mishin, Correlation between grain boundary energy and geometry in Ni-rich NiAl, Acta Materialia 53 (14) (2005) 3795–3805. doi:10.1016/j.actamat.2005.04.043.

[70] H. Beladi, G. S. Rohrer, The relative grain boundary area and energy distributions in a ferritic steel determined from three-dimensional electron backscatter diffraction maps, Acta Materialia 61 (4) (2013) 1404–1412. doi:10.1016/j.actamat.2012.11.017.

[71] X. Zhong, D. J. Rowenhorst, H. Beladi, G. S. Rohrer, The five-parameter grain boundary curvature distribution in an austenitic and ferritic steel, Acta Materialia 123 (2017) 136–145.

[72] V. Randle, G. S. Rohrer, H. M. Miller, M. Coleman, G. T. Owen, Five-parameter grain boundary distribution of commercially grain boundary engineered nickel and copper, Acta Materialia 56 (10) (2008) 2363–2373. doi:10.1016/j.actamat.2008.01.039.

[73] S. Dillon, G. Rohrer, Measuring the Grain Boundary Character and Energy Distributions of Ceramics From Serial Sections of Orientation Maps, Microscopy and Microanalysis 15 (S2) (2009) 608–609. doi:10.1017/s1431927609094148.

[74] G. S. Rohrer, Others, The grain boundary data archive, accessed: 2021-12-03 (2021).
URL http://mimp.materials.cmu.edu/~gr20/Grain_Boundary_Data_Archive/

[75] K. Barmak, J. Kim, C.-S. Kim, W. Archibald, G. Rohrer, A. Rollett, D. Kinderlehrer, S. Ta'asan, H. Zhang, D. Srolovitz, Grain boundary energy and grain growth in Al films: Comparison of experiments and simulations, Scripta Materialia 54 (6) (2006) 1059–1063. doi:10.1016/j.scriptamat.2005.11.060.

[76] G. S. Rohrer, E. A. Holm, A. D. Rollett, S. M. Foiles, J. Li, D. L. Olmsted, Comparing calculated and measured grain boundary energies in nickel, Acta Materialia 58 (15) (2010) 5063–5069. doi:10.1016/j.actamat.2010.05.042.

[77] National Science and Technology Council, Materials Genome Initiative for Global Competitiveness, Tech. rep., US Office of Science and Technology Policy, Washington D.C. (2011).

[78] S. Patala, C. A. Schuh, Symmetries in the representation of grain boundary-plane distributions, Philosophical Magazine 93 (5) (2013) 524 – 573. doi:10.1080/14786435.2012.722700.

[79] E. R. Homer, E. A. Holm, S. M. Foiles, D. L. Olmsted, Trends in grain boundary mobility: Survey of motion mechanisms, JOM-Journal Of The Minerals Metals & Materials Society 66 (1) (2014) 114–120. doi:10.1007/s11837-013-0801-2.

[80] T. Yu, S. Yang, C. Deng, Survey of grain boundary migration and thermal behavior in Ni at low homologous temperatures, Acta Materialia 177 (2019) 151 – 159. doi:10.1016/j.actamat.2019.07.034.

[81] E. R. Homer, S. M. Foiles, E. A. Holm, D. L. Olmsted, Phenomenology of shear-coupled grain boundary motion in symmetric tilt and general grain boundaries, Acta Materialia 61 (4) (2013) 1048–1060. doi:10.1016/j.actamat.2012.10.005.

[82] T. Yu, I. Chesser, S. Ratanaphan, E. Holm, S. Yang, C. Deng, Survey of shear coupling behavior in FCC Ni and BCC Fe grain boundaries, Materialia 15 (2021) 100945. doi:10.1016/j.mtla.2020.100945.

[83] S. Pal, K. V. Reddy, T. Yu, J. Xiao, C. Deng, The spectrum of atomic excess free volume in grain boundaries, Journal of Materials Science 56 (19) (2021) 11511–11528. doi:10.1007/s10853-021-06028-4.

[84] O. K. Johnson, E. R. Homer, D. T. Fullwood, D. E. Page, K. F. Varela, S. G. Baird, Inference and uncertainty propagation of GB structure-property models: H Diffusivity in [100] tilt GBs in Ni, Acta Mater. (215) (2021) 116967. doi:10.1016/j.actamat.2021.116967.

[85] S. G. Baird, E. R. Homer, D. T. Fullwood, O. K. Johnson, Five degree-of-freedom property interpolation of arbitrary grain boundaries via Voronoi fundamental zone framework, Computational Materials Science 200 (2021) 110756. doi:10.1016/j.commatsci.2021.110756.

[86] S. E. Restrepo, S. T. Giraldo, B. J. Thijsse, Using artificial neural networks to predict grain boundary energies, Computational Materials Science 86 (2014) 170–173. doi:10.1016/j.commatsci.2014.01.039.

[87] E. R. Homer, G. L. W. Hart, D. Hensley, C. B. Owens, L. H. Serafin, Computed al grain boundary structures and energies covering 5d space, Mendeley Data, v1 (2022). doi:10.17632/4ykjz4ngwt.

[88] R. Quey, A. Villani, C. Maurice, Nearly uniform sampling of crystal orientations, Journal of Applied Crystallography 51 (4) (2018) 1162–1173. doi:10.1107/s1600576718009019.

[89] D. Brandon, B. Ralph, S. Ranganathan, M. Wald, A field ion microscope study of atomic configuration at grain boundaries, Acta Metallurgica 12 (7) (1964) 813–821. doi:10.1016/0001-6160(64)90175-0.

[90] B. Chalmers, H. Gleiter, A re-interpretation of the 'coincidence model' for grain boundaries, Philosophical Magazine 23 (186) (1971) 1541–1546. doi:10.1080/14786437108217020.

[91] A. P. Sutton, R. W. Balluffi, Overview no. 61 On geometric criteria for low interfacial energy, Acta Metallurgica Et Materialia 35 (9) (1987) 2177–2201. doi:10.1016/0001-6160(87)90067-8.

[92] V. Randle, On the relationship between grain boundary misorientation and grain boundary plane - Response to 'Current issues in recrystallisation: a review', Materials Science And Engineering A-Structural Materials Properties Microstructure And Processing 259 (1) (1999) 138–140.

[93] V. Randle, 'Special' boundaries and grain boundary plane engineering, Scripta Materialia 54 (6) (2006) 1011–1015. doi:10.1016/j.scriptamat.2005.11.050.

[94] F. Otto, E. J. Payton, J. Frenzel, G. Eggeler, The effectiveness of coincidence site lattice criteria in predicting creep cavitation resistance, Journal of Materials Science 47 (6) (2012) 2915–2927. doi:10.1007/s10853-011-6124-1.

[95] E. R. Homer, Investigating the mechanisms of grain boundary migration during recrystallization using molecular dynamics, in: 36th Riso International Symposium on Materials Science, IOP Publishing, Roskilde, 2015, p. 012006. doi:10.1088/1757-899X/89/1/012006.

[96] E. B. Tadmor, R. E. Miller, Modeling Materials, Continuum, Atomistic and Multiscale Techniques, Cambridge University Press, Cambridge, 2011.

[97] Y. Mishin, D. Farkas, M. J. Mehl, D. A. Papaconstantopoulos, Interatomic potentials for monoatomic metals from experimental data and ab initio calculations, Physical Review B 59 (5) (1999) 3393–3407. doi:10.1103/physrevb.59.3393.

[98] A. P. Thompson, H. M. Aktulga, R. Berger, D. S. Bolintineanu, W. M. Brown, P. S. Crozier, g. J. family=in 't Veld, A. Kohlmeyer, S. G. Moore, T. D. Nguyen, R. Shan, M. J. Stevens, J. Tranchida, C. Trott, S. J. Plimpton, LAMMPS - a flexible simulation tool for particle-based materials modeling at the atomic, meso, and continuum scales, Computer Physics Communications 271 (2022) 108171. doi:10.1016/j.cpc.2021.108171.

[99] F. Abdeljawad, S. M. Foiles, A. P. Moore, A. R. Hinkle, C. M. Barr, N. M. Heckman, K. Hattar, B. L. Boyce, The role of the interface stiffness tensor on grain boundary dynamics, Acta Materialia 158 (2018) 440 – 453. doi:10.1016/j.actamat.2018.06.025.

[100] R. D. Moore, T. Beecroft, G. S. Rohrer, C. M. Barr, E. R. Homer, K. Hattar, B. L. Boyce, F. Abdeljawad, The grain boundary stiffness and its impact on equilibrium shapes and boundary migration: Analysis of the $\sigma$ 5, 7, 9, and 11 boundaries in ni, Acta Materialia 218 (2021) 117220. doi:10.1016/j.actamat.2021.117220.

[101] A. Stukowski, Visualization and analysis of atomistic simulation data with OVITO-the Open Visualization Tool, Modelling And Simulation In Materials Science And Engineering 18 (1) (2010) 015012. doi:10.1088/0965-0393/18/1/015012.

[102] F. Musil, A. Grisafi, A. P. Bartók, C. Ortner, G. Csányi, M. Ceriotti, Physics-inspired structural representations for molecules and materials, Chemical Reviews 121 (16) (2021) 9759–9815. doi:10.1021/acs.chemrev.1c00021.

[103] A. P. Bartók, R. Kondor, G. Csányi, On representing chemical environments, Physical Review B 87 (18) (2013) 184115. doi:10.1103/PhysRevB.87.184115.

[104] L. M. Ghiringhelli, J. Vybiral, E. Ahmetcik, R. Ouyang, S. V. Levchenko, C. Draxl, M. Scheffler, Learning physical descriptors for

materials science by compressed sensing, New Journal of Physics 19 (2) (2017) 023017. doi:10.1088/1367-2630/aa57bf.

[105] S. De, A. P. Bartók, G. Csányi, M. Ceriotti, Comparing molecules and solids across structural and alchemical space, Physical Chemistry Chemical Physics 18 (20) (2016) 13754–13769. doi:10.1039/C6CP00415F.

[106] P. M. Larsen, S. Schmidt, J. Schiotz, Robust structural identification via polyhedral template matching, Modelling And Simulation In Materials Science And Engineering 24 (5) (2016) 055007. doi:10.1088/0965-0393/24/5/055007.

[107] E. A. Lazar, VoroTop: Voronoi cell topology visualization and analysis toolkit, Modelling And Simulation In Materials Science And Engineering 26 (1) (2018) 015011. doi:10.1088/1361-651X/aa9a01.

[108] G. Csányi, S. Winfield, J. R. Kermode, A. De Vita, A. Comisso, N. Bernstein, M. C. Payne, Expressive programming for computational physics in fortran 95+, IoP Comput. Phys. Newsletter (2007) Spring 2007.

[109] A. P. Bartók, M. C. Payne, R. Kondor, G. Csányi, Gaussian approximation potentials: the accuracy of quantum mechanics, without the electrons, Phys. Rev. Lett. 104 (13) (2010) 136403. doi:10.1103/PhysRevLett.104.136403.

[110] J. R. Kermode, f90wrap: an automated tool for constructing deep python interfaces to modern fortran codes, J. Phys. Condens. Matterdoi:10.1088/1361-648X/ab82d2.

[111] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python, Journal of Machine Learning Research 12 (2011) 2825–2830.

[112] F. Ercolessi, J. B. Adams, Interatomic potentials from first-principles calculations: The force-matching method, Europhysics Letters (EPL) 26 (8) (1994) 583–588. doi:10.1209/0295-5075/26/8/005.

[113] D. Wolf, A read-shockley model for high-angle grain boundaries, Scripta Metallurgica Et Materialia 23 (10) (1989) 1713–1718. doi:10.1016/0036-9748(89)90348-7.

[114] X. W. Zhou, R. B. Sills, D. K. Ward, R. A. Karnesky, Atomistic calculations of dislocation core energy in aluminium, Physical Review B 95 (5) (2017) 054112. doi:10.1103/physrevb.95.054112.

[115] R. Sarochawikasit, C. Wang, P. Kumam, H. Beladi, T. Okita, G. S. Rohrer, S. Ratanaphan, Grain boundary energy function for $\alpha$ iron, Materialia 19 (2021) 101186. doi:10.1016/j.mtla.2021.101186.

[116] O. Chirayutthanasak, R. Sarochawikasit, A. Wisitsorasak, N. Rujisamphan, T. Frolov, T. Oppelstrup, S. Dangtip, G. S. Rohrer, S. Ratanaphan, Anisotropic grain boundary area and energy distributions in tungsten, Scripta Materialia 209 (2022) 114384. doi:10.1016/j.scriptamat.2021.114384.

[117] J. K. Mackenzie, M. J. Thomson, Some Statistics Associated with the Random Disorientation of Cubes, Biometrika 44 (1/2) (1957) 205. doi:10.2307/2333253.

[118] J. K. Mackenzie, Second Paper on Statistics Associated with the Random Disorientation of Cubes, Biometrika 45 (1/2) (1958) 229. doi:10.2307/2333059.

[119] V. Randle, Twinning-related grain boundary engineering, Acta Materialia 52 (14) (2004) 4067–4081. doi:10.1016/j.actamat.2004.05.031.

[120] W. Wang, C. Cai, G. S. Rohrer, X. Gu, Y. Lin, S. Chen, P. Dai, Grain boundary inter-connections in polycrystalline aluminum with random orientation, Materials Characterization 144 (2018) 411–423. doi:10.1016/j.matchar.2018.07.040.

[121] X. Zhu, R. Birringer, U. Herr, H. Gleiter, X-ray diffraction studies of the structure of nanometer-sized crystalline materials, Physical Review B 35 (17) (1986) 9085–9090. doi:10.1103/physrevb.35.9085.

[122] W. Wunderlich, Y. Ishida, R. Maurer, HREM-studies of the microstructure of nanocrystalline palladium, Scripta Metallurgica et Materialia 24 (2) (1990) 403–408. doi:10.1016/0956-716x(90)90279-p.

[123] H. Kuwano, H. Ouyang, B. Fultz, A Mössbauer spectrometry study of nanophase Cr-Fe synthesized by mechanical alloying: A measurement of grain boundary width, Nanostructured Materials 1 (2) (1992) 143–148. doi:10.1016/0965-9773(92)90067-8.

[124] D. E. Page, K. F. Varela, O. K. Johnson, D. T. Fullwood, E. R. Homer, Measuring simulated hydrogen diffusion in symmetric tilt nickel grain boundaries and examining the relevance of the Borisov relationship for individual boundary diffusion, Acta Materialia 212 (2021) 116882. doi:10.1016/j.actamat.2021.116882.

[125] D. M. Saylor, A. Morawiec, G. S. Rohrer, Distribution of grain boundaries in magnesia as a function of five macroscopic parameters, Acta Materialia 51 (13) (2003) 3663–3674. doi:10.1016/S1359-6454(03)00181-2.

[126] S. G. Baird, E. R. Homer, D. T. Fullwood, O. K. Johnson, Towards a quantitative cartography of the grain boundary energy landscape: Paths and correlations, ChemRxivdoi:10.26434/chemrxiv-2021-ds0ml.

[127] T. Francis, I. Chesser, S. Singh, E. A. Holm, M. D. Graef, A geodesic octonion metric for grain boundaries, Acta Materialia 166 (2019) 135 – 147. doi:10.1016/j.actamat.2018.12.034.

[128] I. Chesser, T. Francis, M. D. Graef, E. A. Holm, Learning the grain boundary manifold: tools for visualizing and fitting grain boundary properties, Acta Materialia 195 (2020) 209 – 218. doi:10.1016/j.actamat.2020.05.024.

[129] C. Hu, Y. Zuo, C. Chen, S. P. Ong, J. Luo, Genetic algorithm-guided deep learning of grain boundary diagrams: Addressing the challenge of five degrees of freedom, Materials today 38 (2020) 49 – 57. doi:10.1016/j.mattod.2020.03.004.

[130] J. K. Mason, S. Patala, Basis functions on the grain boundary space: Theory, arXivarXiv:1909.11838.

[131] D. Foley, G. J. Tucker, Quantifying grain boundary damage tolerance with atomistic simulations, Modelling and Simulation in Materials Science and Engineering 24 (7) (2016) 075011. doi:10.1088/0965-0393/24/7/075011.

[132] A. R. Oganov, C. W. Glass, Crystal structure prediction using ab initio evolutionary techniques: Principles and applications, The Journal of Chemical Physics 124 (24) (2006) 244704. doi:10.1063/1.2210932.

[133] A. O. Lyakhov, A. R. Oganov, H. T. Stokes, Q. Zhu, New developments in evolutionary structure prediction algorithm USPEX, Computer Physics Communications 184 (4) (2013) 1172–1182. doi:10.1016/j.cpc.2012.12.009.

[134] K. C. Alexander, C. A. Schuh, Exploring grain boundary energy landscapes with the activation-relaxation technique, Scripta Materialia 68 (12) (2013) 937–940. doi:10.1016/j.scriptamat.2013.02.038.

[135] K. C. Alexander, C. A. Schuh, Towards the reliable calculation of residence time for off-lattice kinetic Monte Carlo simulations, Modelling And Simulation In Materials Science And Engineering 24 (6) (2016) 065014. doi:10.1088/0965-0393/24/6/065014.

[136] I. S. Winter, R. E. Rudd, T. Oppelstrup, T. Frolov, Nucleation of Grain Boundary Phases, Physical Review Letters 128 (3) (2022) 035701. doi:10.1103/physrevlett.128.035701.

[137] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner, Gradient-based learning applied to document recognition, Proceedings of the IEEE 86 (11) (1998) 2278–2324.

[138] L. Deng, The MNIST Database of Handwritten Digit Images for Machine Learning Research [Best of the Web], IEEE Signal Processing Magazine 29 (6) (2012) 141–142. doi:10.1109/msp.2012.2211477.

[139] Y. Bengio, A. Courville, P. Vincent, Representation Learning: A Review and New Perspectives, IEEE Transactions on Pattern Analysis and Machine Intelligence 35 (8) (2013) 1798–1828. doi:10.1109/tpami.2013.50.