Computer Physics Communications 232 (2018) 256–263

Contents lists available at ScienceDirect

Computer Physics Communications

journal homepage: www.elsevier.com/locate/cpc

Imeall: A computational framework for the calculation of the atomistic properties of grain boundaries☆

H. Lambert [a,*], Adam Fekete [a], J.R. Kermode [b], A. De Vita [a,c]

[a] Department of Physics, King's College London, Strand, London, WC2R 2LS, United Kingdom
[b] Warwick Centre for Predictive Modelling, School of Engineering, University of Warwick, Coventry CV4 7AL, United Kingdom
[c] Department of Engineering and Architecture, University of Trieste, I-34127 Trieste, Italy

ARTICLE INFO

Article history:
Received 29 September 2017
Received in revised form 31 March 2018
Accepted 21 April 2018
Available online 7 May 2018

Keywords:
Grain boundaries
Interatomic potentials
Atomistic simulation
Dislocation analysis
Database
Hydrogen embrittlement

ABSTRACT

We describe the Imeall package for the calculation and indexing of atomistic properties of grain boundaries in materials. The package provides a structured database for the storage of atomistic structures and their associated properties, equipped with a programmable application interface to interatomic potential calculators. The database adopts a general indexing system that allows storing arbitrary grain boundary structures for any crystalline material. The usefulness of the Imeall package is demonstrated by computing, storing, and analysing relaxed grain boundary structures for a dense range of low index orientation axis symmetric tilt and twist boundaries in α-iron for various interatomic potentials. The package's capabilities are further demonstrated by carrying out automated structure generation, dislocation analysis, interstitial site detection, and impurity segregation energies across the grain boundary range. All computed atomistic properties are exposed via a web framework, providing open access to the grain boundary repository and the analytic tools suite.

Program summary
Program Title: Imeall
Program Files doi: http://dx.doi.org/10.17632/nj77stc62b.1
Licensing provisions: Apache-2.0
Programming languages: python, fortran, javascript
Nature of problem: Determining the minimum energy structure for a specific grain boundary and interatomic force field involves extensive searches in configuration space. Duplication of this effort should be avoided, and providing a unique database to gather the resulting structures is needed for this. Accurately cataloguing the chemical and electronic environments associated with the interface atoms in the grain boundary furthermore requires a useable interface between a database of grain boundary structures and an expandible set of interatomic potential calculators.
Solution method: We introduce a standard indexing convention that allows the integration of grain boundary structures for arbitrary materials, generated by different users/research groups, into a normalised database that can be easily queried and used as a starting point for further research projects. The Imeall package accomplishes this by specifying a standard naming convention for the grain boundary database and by providing the software routines necessary to populate and query such a database, as well as an interface to interatomic calculators and analysis tools.
References:
[1] QUIPPY https://libatoms.github.io/QUIP/
[2] ASE https://wiki.fysik.dtu.dk/ase/
[3] OVITO https://www.ovito.org
[4] Transformations http://www.lfd.uci.edu/~gohlke/code/transformations.py
[5] pylada-defect https://github.com/pylada/pylada-defects

1. Introduction

A grain boundary is formed wherever an extended planar region can be identified that separates two crystalline regions of a homophase material differing in the relative orientation of the adjoining crystal lattices. Grain boundaries play a significant role in

☆ This paper and its associated computer program are available via the Computer Physics Communication homepage on ScienceDirect (http://www.sciencedirect.com/science/journal/00104655).

* Corresponding author.
E-mail address: henry.lambert@kcl.ac.uk (H. Lambert).

https://doi.org/10.1016/j.cpc.2018.04.029

determining the mechanical properties of materials. Having access to a standardised database of grain boundary structures is an important prerequisite for any atomic-scale ("chemo-mechanical") analysis, e.g., to investigate how impurities trapped at grain boundaries influence the mechanical strength of the material. Evaluating several properties of significant engineering interest requires access to the atomistic structures of grain boundaries. These include, but are by no means limited to, the diffusivity of impurities at grain boundary interfaces [1], the segregation energies of impurities to interfaces [2], and the interaction and slip transmission of dislocations across boundaries [3,4]. While a number of systematic investigations of grain boundary structures have been performed, see Refs. [5–7] and references therein, a single repository equipped with the appropriate tools to generate, archive and analyse the boundaries is still missing. The `Imeall` package and the suite of routines described in this work make a step in this direction by providing a framework for constructing and cataloguing grain boundary structures, determining their minimum energy, and calculating an array of quantities of interest.

The paper is organised as follows. Section 2 describes the procedure and routines for systematically generating tilt and twist grain boundaries. Section 3 introduces a unique naming convention which is reflected in the structure of the `Imeall` repository. This scheme allows the database to incorporate grain boundary structures for any reference crystalline material, computed using any interatomic potential, in a consistent and physically intuitive fashion. The information contained in the database is mirrored by a normalised SQL database to allow rapid queries via the command line or a web interface. Section 4 describes the overall layout of the computational package, highlighting the key routines needed to perform the atomistic calculations and analysis, and those used for storing this information in the database and exposing it via a web framework. The following sections describe some applications of the package, to illustrate the comparative information that can be exposed. In particular, Section 5 uses the package to investigate how grain boundary energetics and structural topologies vary upon using different interatomic potential parameterisations. Finally, Section 6 describes how the package handles calculations addressing point-defect energetics, specifically interstitial hydrogen located in the vicinity of α-Fe grain boundaries.

2. Generating grain boundary structures

In most cases the full macroscopic specification of a grain boundary can be accomplished using five degrees of freedom [8].[1] These 5 degrees of freedom constitute thermodynamic variables for the system and specify the boundary conditions far removed from the boundary [8]. To generate a boundary, two identical crystals, which we may here differentiate by referring to them as the 'red crystal' and the 'blue crystal' (see Fig. 1) are initialised in the same orientation. The red crystal is then rotated by an angle θ (a single degree of freedom) around a rotation axis specified by the vector $\hat{N}$ (associated with two more degrees of freedom). After this rotation, a boundary plane is chosen, specified by its normal vector $\hat{bp}$ expressed in the unrotated coordinate system. This final vector exhausts the remaining two degrees of freedom and concludes the macroscopic specification of the grain boundary. In the preceding both $\hat{N}$ and $\hat{bp}$ are unit vectors obtained by normalising the lattice vectors chosen to effect the necessary rotations. All red crystal atoms located below the boundary plane and all blue crystal atoms located above the boundary plane are at this point removed. We will refer to the choice of orientation axis, misorientation angle,

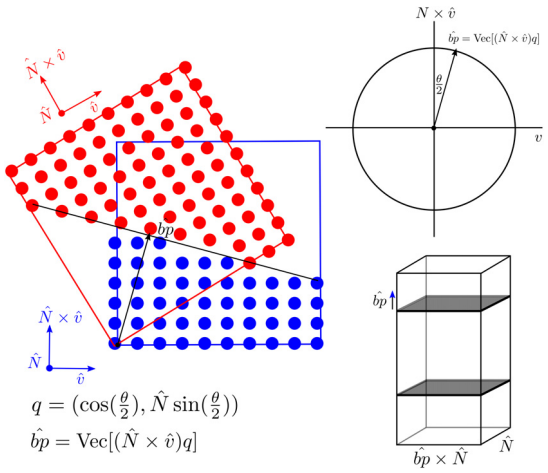

Fig. 1. The general coordinate system for determining tilt and twist boundary planes and orientations using quaternion algebra. Hats indicate regular three dimensional vectors, considered to be quaternions with a 0 component for the scalar part. The quantity q, without a hat, represents a quaternion with a vector part determined by the orientation axis $\hat{N}$ and a scalar part determined by the misorientation angle θ. The effects of quaternion multiplication on the vectors normal to the orientation axis plane are represented schematically. Upper right inset: the total relative rotation θ of the two misoriented grains is depicted by the red and blue crystals with their coordinate systems rotated. Lower right inset: the bicrystal generated by reflecting the grain boundary to double the number of grain boundaries in a unit cell. This guarantees periodicity perpendicular to the grain boundary plane allowing for the use of periodic boundary conditions in the atomistic relaxation.

and boundary plane as a complete *canonical* macroscopic specification of the grain boundary, i.e., before atomistic relaxations are considered. It is useful for what follows to define two grain boundary main geometries. If the rotation axis is parallel to the boundary plane normal ($\hat{N}$ parallel to $\hat{bp}$), the grain boundary is referred to as a twist boundary. If the rotation axis is orthogonal to the boundary plane normal ($\hat{N}$ orthogonal to $\hat{bp}$), the grain boundary is referred to as a tilt boundary. In the case where the Miller indices of the boundary plane are the same in the coordinate systems of both grains, the grain boundary is referred to as symmetric. In the `Imeall` package generating an exhaustive array of tilt symmetric and twist grain boundary structures is accomplished by the methods defined in `imeall.slabmaker.slabmaker` and `imeall.slabmaker.gengb_from_quat`.

These routines make use of quaternion algebra to systematically generate a full array of symmetric tilt and twist boundaries (see Refs. [9,10] for other applications of quaternion algebra to the study of grain boundaries; an overview of quaternion algebra is given in Ref. [11]).

Quaternions are frequently used in engineering and graphic design contexts to handle vector rotations. For the present purposes, quaternions can be thought of as four dimensional objects with one scalar component and a three-dimensional vector component. The rotation of a vector $\hat{v}$ by an angle θ around a unit vector $\hat{N}$ is accomplished by conjugation of $\hat{v}$ by the quaternion q, i.e., by multiplying the vector on the left and the right as $q\hat{v}q^{-1}$, where q^{-1} denotes the inverse of q. In the context of grain boundaries the four components of the quaternion, $q = (\theta, \hat{N})$ are converted to $q = (\cos(\theta/2), \sin(\theta/2)\hat{N})$ so that the necessary rotations may be obtained. These components are directly obtained from the physical parameters defining the grain boundary, i.e. the misorientation angle and the orientation axis. Interestingly, if the quaternion q can be reduced to a *primitive form*, i.e. all its entries are integers apart from a scaling factor, the rotation is guaranteed to produce

[1] In the case of enantiomorphic crystals the chirality of the crystal at the interface can constitute a 6th macroscopic degree of freedom.

a coincident site sublattice (cf. Refs. [12,13]) that is, a periodic sublattice of space points where red and blue atoms coincide. In this case, given q, closed expressions exist that readily provide a set of basis vectors for the coincident site sublattice [10].

Besides making it easy to perform rotations with no need for using rotation matrices, there is an additional favourable aspect of the quaternion scheme for generating symmetric tilt grain boundaries. This is illustrated in Fig. 1, which presents a schematic representation for the determination of the boundary planes for a tilt symmetric grain boundary using quaternion algebra. The two crystals, red and blue, are initially superimposed. In Fig. 1 the unit vector denoting the axis where the misorientation angle is measured from, according to a right hand rule, is denoted $\hat{N} \times \hat{v}$. Here we have assumed that $\hat{N}$ is normal to a lattice plane (the coordinates of $\hat{N}$ will be thus three integers apart from a scaling factor) and $\hat{v}$ is a lattice vector of the plane, orthogonal to $\hat{N}$. The right quaternion ("half rotation") product $(N \times v)q$ can be used in this geometry to define the boundary plane normal which in the original reference frame corresponds to the symmetric tilt grain boundary plane, i.e., the grain boundary associated with the misorientation angle θ and axis $\hat{N}$ encoded as components of q. All red lattice vectors $\hat{\tilde{v}}$ can be obtained at this point by a left and right quaternion product ("full rotation"), as $q\hat{\tilde{v}}q^{-1}$. Once all atomic positions of the grain boundary interface have been generated, the unit cell is doubled by reflecting it through the boundary plane, to produce a periodic structure (cf. Fig. 1). In the `Imeall` package the `gen_sym_tilt` method generates an array of angle boundary plane pairs, and the `build_tilt_sym_gb` generates the bicrystal unit cells according to these specifications. Equivalent routines exist for twist boundary structures.

2.1. Microscopic search parameters

While the macroscopic degrees of freedom of the grain boundary are completely specified by the geometric considerations, at the interface itself there remain four microscopic degrees of freedom that must adjust to satisfy the macroscopic boundary conditions. The four microscopic degrees of freedom consist of in plane rigid body translations (2 degrees of freedom), expansion normal to the boundary plane (the third degree of freedom), and for non-monoatomic crystals the position of the boundary plane. This search requires initialising the misoriented grains in a reference frame defined by a two dimensional grid of unique positions, $\hat{x}$, $\hat{y}$, the rigid body translations, representing the in-plane translation vectors, and a lattice expansion vector $\hat{e}$, normal to the boundary plane, which allows for atomistic relaxation of the initial grain boundary structure. For further details regarding the requirements of performing this search see Ref. [14] and Refs. [7,15–17]. The mechanics of carrying out the grain boundary structural relaxation search and the routines required for performing the atomistic relaxations are detailed in Sections 3 and 5. The methods needed to generate the atomistic grain boundary bicrystal and the graphical representation of coincident site lattices, are handled in the `imeall.slabmaker.slabmaker` module. All computations above defining the atomistic cells are performed using the `ase` library. The resulting structures are stored as extended `.xyz` files in the grain boundary database. The extended `.xyz` is a file format that allows for the specification of the atomic structures in plain text format with the number of atoms, their type, the cell geometry and cartesian coordinates comprising the minimum information required. Extended `.xyz` files also allow per atom properties to be stored: these properties can be vectors, tensors or scalar properties to carry information, e.g., forces or local magnetic moments.

3. Grain boundary database

3.1. Grain boundary hierarchy

In the `Imeall` package each grain boundary is assigned an identifying label, an "id" tag, which is built up in an intentional analogy with the physical specification of the grain boundary. Due to the large number of additional conventions that become necessary when discussing non-cubic systems we have elected not to use the "mean boundary plane" formalism for specifying the grain orientation [8].

Rather, we choose an alternative specification that arises naturally from the procedure described in the previous sections, which involves initialising two superimposed crystals oriented in the form of a conventional unit cell, applying rotations via a right hand rule, and following an intuitive geometric procedure for initialising the configuration. For this, the orientation axis $\hat{N}$, the angle θ (in degrees), and the boundary plane normal vector $\hat{bp}$ are serialised as absolute value integers and concatenated into a string with five digits reserved for the angle and negative integers prepended with the character 'm', and the geometric degrees of freedom separated by an underscore delimiter. For instance, a symmetric tilt boundary specified by the orientation axis [110], with a misorientation angle of 13.44 degrees and separated by the $(\bar{1}\ 1\ 12)$ boundary plane is identified by the grain boundary id `1_1_0_01344_m1_1_12`. The grain boundary at this description level is referred to as the 'canonical grain'. The database is constructed by adopting a nested hierarchy starting from the common root position, which is followed by the material directory, the orientation axis directory, and finally the canonical grain id. As an example, assuming that the material is α-Fe and the database root is stored at '`/`', the canonical grain mentioned above would be accessed at '`/alphaFe/1_1_0/1_1_0_01344_m1_1_12`'.[2]

To complete the canonical grain descriptor level (at which point there is still no assumptions about the interatomic potential used for detailed atomistic characterisation), any data determined purely by geometric considerations, e.g. the orientation axis or the grain boundary Σ number (i.e. the inverse ratio of the number of coincident sites to lattice sites), is included in a file called `gb.json`. The unrelaxed canonical grain is at the same time stored as an extended `.xyz` file, along with an image of the grain, and a schematic vector graphics image of two planes of the coincident site lattice at the interface.

Below the canonical grain descriptor level, the hierarchy proceeds in a general way intended to capture the full range of accessible microscopic structures. For this, the output of a series of subsequent microscopic initialisations–relaxations procedures, set up according to the requirements described in Section 2.1, is stored in a unique `Potential` directory.

The hierarchy at this level is flat with each interatomic potential occupying a separate directory. The starting point for relaxing initialised structures could be a computationally inexpensive calculation performed using an EAM or GAP potential, or an empirical tight binding model, followed by refining with various levels of DFT.

The most interesting structures from the relaxation stage performed with the less computationally expensive calculators would then be forwarded to the more stringently parameterised calculator in the next `Potential` directory; e.g., `EAM_1`, `GAP_1`, `DFT_1`, and `DFT_2`. The usefulness of this procedure is demonstrated by its

[2] For a heterophase interface the top level of the canonical name would be generated by concatenating the string specifying the two materials, e.g., α-Fe–Fe_3C for a BCC iron–cementite interface. Subsequent levels of the naming hierarchy would have to be similarly doubled to reflect the orientation axes, angle, and boundary planes in the respective crystals.

Table 1
The directory naming convention in `Imeall` with an example grain. The example column has been chosen to uniquely specify the minimum energy structure for the [1 1 0] 13.44 $(\bar{1}112)$ canonical grain boundary. The chosen interatomic potential is of the embedded atom type and is specified in Ref. [20]. The pathname for this grain boundary in the directory structure would then be `1_1_0_01344_m1_1_12_v6bxv2_tv0.4bxv0.2_d1.4z` denoting a 6×2 supercell with in plane rigid body translations of 0.4, 0.2 (referring to the fraction translation in the basis of the lattice vectors of the primitive unit cell for an orthorhombic grain boundary), and an atom deletion criterion for nearest neighbours below 1.4 Å.

Specification	Example	String
Orientation axis	[1 1 0]	1_1_0
Angle	13.44	01344
Boundary plane	$(\bar{1}\ 1\ 12)$	m1_1_12
Supercell	6×2	_v6bxv2
Rigid body translation	[0.4, 0.2]	_tv0.4bxv0.2
Atom deletion criterion	1.4Å	_d1.4z

ability to handle issues previously identified in Refs. [18,19] where classical potentials identify non-physical minimum energy grain boundary structures. The minimum energy structure identified by embedded atom potentials has a tendency to favour structures preserving the coordination number at the interface. This propensity for over-bonding introduces the risk of creating structures with an unrealistic angular distortion. Quantum mechanical models can, however, restore the physically correct angular contributions to the total energy.

The separation by potential immediately below the canonical level is emphasised so that the predicted atomistic structures for different potentials remain distinct. It also facilitates the iterative generation of accurate atomistic models of interfaces, and allows for direct comparison of predicted structures with experimentally observed interfacial structures.

Continuing down the hierarchy, the different microscopic structural initialisations–relaxations can be constructed and labelled systematically within a subdirectory structure branching underneath each potential directory: see Table 1 for a description of the labelling convention and an example grain specification. The appropriate number of microscopic initialisations–relaxations to be stored will depend on the application.

Constructing the database in the described manner facilitates the direct tabulation and comparison of quantities for a definite structure with an exact geometric specification, i.e. a grain boundary, as predicted by different interatomic potential models of arbitrary accuracy. The full microscopic initialisation is determined by the supercell size of the subgrain and the rigid body translations. For multi-atom systems and heterophase interfaces additional information on the microscopic positioning of the interfacial plane can also be appended in the naming convention but this is left for future iterations of the database. The final microscopic degree of freedom that must be varied is the atom deletion criterion distance. The atom deletion criterion arises because of the way grain boundary structures are generate. Computationally the necessary rotations of the crystals are performed to initialise the bicrystal structure. This procedure can result in atoms placed unphysically close together. To avoid missing initial configurations that might result in valid interfacial structures upon relaxation, the distance criterion for removing neighbouring atoms that are below a certain threshold is varied systematically. For Fe grain boundaries the lower end of the threshold is chosen around 1.4 Å and stepped up to 2.5 Å in steps of 0.1 Å. This criterion can be chosen to be as fine grained as necessary. Depending on the orientation of the two grains and the in plane lattice translations chosen there will be a spectrum of atomic nearest neighbour distances at the interface. The final grain boundary structure arrived at will be dependent on the initial choice of this deletion criterion.

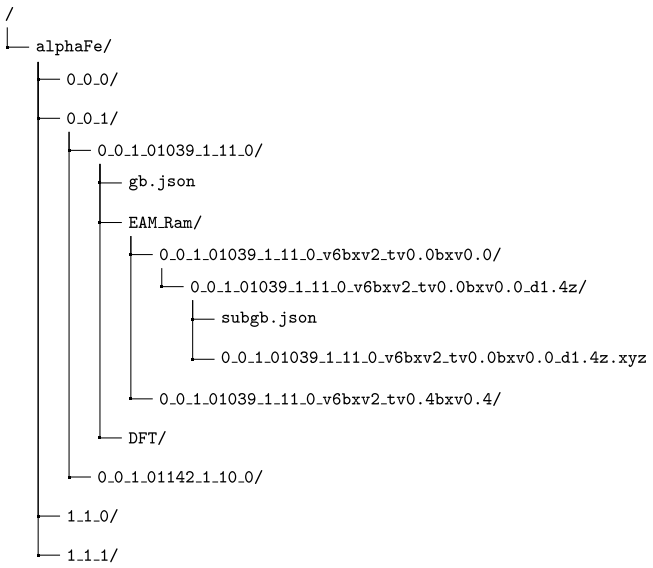

Fig. 2. Snapshot of an `Imeall` directory tree. In this case boundaries for α-Fe with the [001], [110], [111] orientation axis exist. Underneath the [001] orientation axis we show the layout for the [001] 10.39 [1 11 0] grain boundary. Two interatomic potentials have been used: the EAM parameterised in Ref. [20] and a density functional theory calculation (with precise parameterisations to be included in the dedicated `.json` file. Below the EAM directory a specific `SubGrainBoundary`, according to the naming convention in Table 1, is represented.

The combination of a given potential and fully specified subgrain structure is referred to as a `SubGrainBoundary`. The predicted atomistic structures of a grain boundary will inevitably vary, depending on the interatomic potential model used in the calculation. Organising the database of microscopic structures according to potential models implies that the database can seamlessly accommodate newly predicted structures for any grain boundary, contributed from researchers using different interatomic potentials. The database also makes provisions for some extra flexibility not needed for standard grain boundaries. In particular, the 'orientation axis' [000] is reserved for isolated dislocations and fracture geometries. The equivalent of the canonical grain in this case specifies the type of dislocation, e.g. edge (e), screw (s), and mixed dislocation (d), or the fracture geometry (f). In the case of an isolated dislocation, the Burgers vector and the dislocation line in serialised form follow the alphabetic specification. For example, an isolated edge dislocation line oriented along the [100] line with a Burgers vector pointing along [010] would be specified as `e_1_0_0_0_1_0`. As an example of a fracture geometry, `f_1_1_1_1_1_m2`, specifies a fracture with the cleavage plane (111) and crack front oriented along the [1 1 −2] direction. Again, a `gb.json` file specifies the essential parameters of the canonical dislocation with subgrain database entries being resolved according to the interatomic potential used (see Fig. 2).

3.2. Closure tree

While a hierarchical database layout is a natural choice for organising our microscopic structure database, executing repeated recursive searches of the directory tree to extract information can become computationally expensive. For this reason, the grain boundary directory hierarchy is also 'flattened' within `Imeall`, using a closure tree, that stores all the intermediate paths to grain boundary structures, which can be rapidly queried. This mirrored database tracks the physical database layout and ensures the grain boundary database remains properly normalised: i.e. there is no duplication of canonical grain structures, or SubGrainBoundary

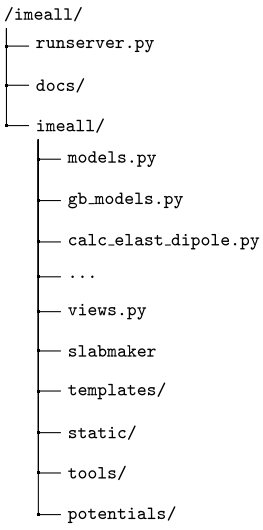

Fig. 3. The core directory structure and routines of the `Imeall` package. The directories `static` and `templates`contain the javascript and html templates for the web interface, the `potentials` directory contains the parameterisation files for the interatomic potentials present in the database. The `models.py` and `gb_models.py` routines define the methods and schema for traversing the grain boundary directory tree and synchronising the directory database and the SQL database. These routines provide the skeleton of the database. Routines prefixed `calc_...` contain logic for various types of analysis, e.g. calculating elastic dipole tensors, performing atomistic relaxations, probing interstitial energetics, etc. The `slabmaker` module contains routines for initialising grain boundary structures.

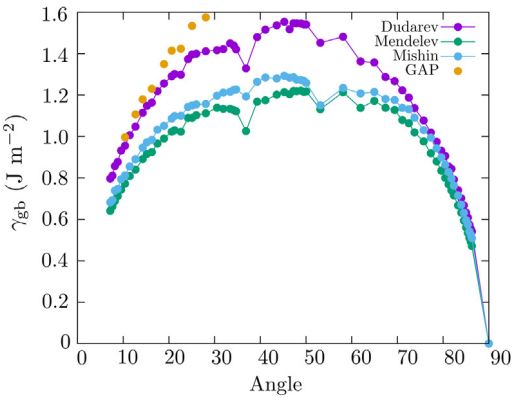

Fig. 4. Variation of grain boundary energies with misorientation angle for the [001] orientation axis for a variety of embedded atom potentials and a GAP interatomic potential model. The respective references are Mendelev [22], Dudarev [23], Mishin [24], and using the Gaussian Approximation Potential [25,26]. The database contains the energetics for each potential along with the configuration space searched to obtain the energetic minimum structure. The GAP model is only applied to the low angle boundaries where there are distinct dislocation networks.

objects for a given potential and there is minimal redundancy in the storage of grain boundary properties. Integrity keys in the database ensure a unique entry for a given potential and a given grain boundary id. This serialised database can also be queried rapidly without having to resort to a recursive search strategy. The `imeall.gb_models` and `imeall.models` modules define the database schema for the SQLite database and provides a number of objects and methods for rapidly querying and retrieving data. It is through this serialised database that the web framework for the `Imeall` package retrieves data for the connected user. Using this bimodal approach to storage allows the database to coexist in the form of an intuitive physical hierarchical layout where a researcher can manually extend and work within a directory tree, and a complimentary serialised database which can be queried in a structured and time-efficient manner.

4. Package layout

The `Imeall` package follows the standard template for a python package. Fig. 3 demonstrates the directory structure of the package and the key routines contained as *.py* files. The `Imeall` package can operate in a distributed or a local mode. In the distributed case the host server can be queried and precomputed structures can be checked out or inspected via the web interface. Alternatively the entire `Imeall` package can be downloaded and installed by any user on a local machine, and, by running the `runserver.py` script a local instance of the server can be used that connects to whatever portion of the grain boundary database the user chooses to store locally.

5. Energetics of relaxed grain boundary structures

As a first application of the package, we describe the automated procedure for generating and relaxing a comprehensive array of grain boundary microscopic configurations. The generation of these job arrays is handled by `run_gb_net.py`. This script generates a set of initial microscopic structures for a given canonical grain boundary geometry, as required to approximately span all microscopic initialisations at the boundary. It then calls the desired atomistic calculator in `ase` to handle the structural minimisations.[3] If the desired interatomic potential is not supported in the `ase` package, the routines in `Imeall` will still generate the structures, which can then be manually forwarded to the calculator of choice. The relaxation method is customisable with the default set to be the FIRE algorithm of Ref. [21]. The structural relaxation proceeds until a desired maximum residual force tolerance on the atoms is achieved. A constrained structural minimisation is performed so that only atomic positions and the lattice vector orthogonal to the grain boundary plane can vary during the relaxation.[4] This constraint can be imposed on the molecular dynamics using a strain filter made available by the `ase` `Constraints` class. The strain filter is a constraint imposed on the dynamical relaxation which allows the lattice vector orthogonal to the interface to adjust during the course of the relaxation. This choice allows the structural minimisation to be carried out at zero pressure with the orthogonal expansion of the boundary plane constituting a microscopic degree of freedom. Fixing the lattice vector orthogonal to the grain boundary plane would be commensurate to carrying out the relaxation under the conditions of constant volume rather than constant pressure.

Fig. 4 reports the grain boundary energies for the minimum energy structures obtained for the [001] orientation axis and a broad range of misorientation angles, using four different interatomic potentials. The interfacial energies are determined by a combination of the elastic strain energy and the chemical reconstructions that take place [27]. In the low angle regime there is a clear distinction between the two contributions, corresponding to a picture of sections of coherent crystal populated by an array of sparse, parallel dislocation cores. As the dislocation core spacing approaches the lattice plane spacing, the decomposition into separated elastic and chemical contributions to the energy becomes more problematic. Fig. 5 illustrates the relaxed dislocation core structures associated

[3] In order to use the Fe–H EAM potential of Ref. [20] the `quippy` package must be used. Otherwise any atomistic calculator available in the `ase` framework can be used to relax the grain boundary structures.

[4] Allowing variation of the boundary in-plane lattice vectors would amount to applying a long range strain on the regions of "good crystal" away from the interface, resulting in a divergence of the system's energy.

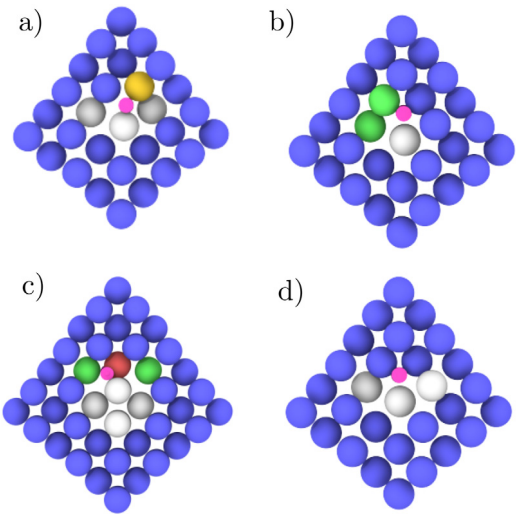

Fig. 5. Minimum energy dislocation cores for a grain boundary formed by a misorientation of 8.17 degrees around the [0 0 1] axis, with a (1 14 0) boundary plane (internal imeall id `0_0_1_00817_1_14_0`). Structures have been computed using a variety of potentials. The colouring reflects the bond angle analysis of Ref. [28] to determine the local atomic environment. Namely, blue corresponds to body centred cubic, red to hexagonal close packed, green to face centred cubic, and gold to icosahedral coordination. Each potential predicts the same spacing between dislocation cores, but there is a significant variety in the local atomic environment predicted by these potentials at each dislocation core. This is accompanied by a corresponding variety of the geometry of interstitial trap sites, which is relevant for point defect diffusivity. Structures a, b, c have been generated using potentials from Refs. [20,23,24], respectively, and structure d was determined using a Gaussian approximation potential [25,26,29].

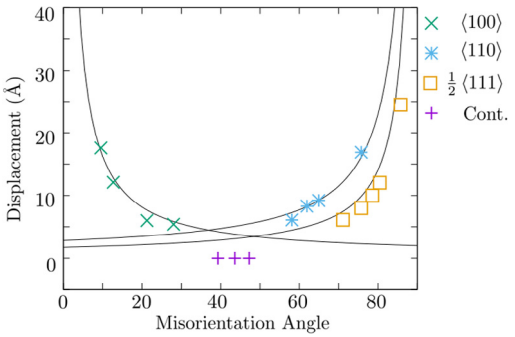

Fig. 6. Dislocation spacing in selected [001] grain boundaries, with the unique Burger's vector as determined by the dislocation extraction algorithm (DXA) (see text). The solid lines are determined by Frank's formula, the marked points correspond to the dislocation spacing determined by atomistic relaxation, and the points marked 'Cont.' refer to large angle grain boundaries where the dislocation cores merge into a continuous network.

with the same grain boundary using four different potentials. The local atomic environment of the atoms at the dislocation core are differentiated using the method of Ref. [28]. The significant structural variation obtained from the different potentials significantly affects the predicted properties of the grain boundary. This motivates the construction of the present single repository as a prerequisite to rationalise the differences.

The equilibrium spacing between dislocations also has a significant effect on the elastic properties of the interface. In low angle grain boundaries the dislocation spacings are governed by Frank's formula:

$$d = \frac{|\mathbf{b}|}{2\sin(\Delta\theta/2)}, \tag{1}$$

where $\mathbf{b}$ is the Burgers vector and the angle θ is measured from the nearest coincident site lattice. Frank's formula readily allows to check whether or not the atomic potential is correctly describing the long range elastic strain field and the Burgers vector of isolated dislocations. The computed dislocation spacings and Burgers vectors are compared with Frank's formula in Fig. 6. The dislocation character and spacing are determined using the dislocation extraction algorithm (DXA) technique developed in Refs. [30,31]. The `Imeall` package also allows for the calculation of the Nye tensor using the technique described in Ref. [32] to identify isolated dislocations. Knowledge of the structure of the dislocations and their spacing allows the construction of analytical models describing the elastic properties of interfaces. The accessibility of such data for a range of grain boundaries and force models provided further motivation for developing the `Imeall` database and tools.

6. Interstitial sites and segregation energies

The possibility of cataloguing trapping sites and trap depths for interstitials is a prerequisite for calculating the diffusivity and equilibrium concentrations of point defects in a material: of particular interest for iron-based materials is hydrogen diffusivity [33], notably in relation to the steel-embrittlement problem. A number of diffusivity models require parameterisations reliant on knowing trapping and segregation energies for boundaries [34–37]. Frequently the distribution of trap site energies is taken to be Gaussian with the variance and centre value of the Gaussian fit to experimental data [35] but this so far had to be assumed rather than calculated with appropriate configuration space statistics.

The `Imeall` database is equipped with the capability of cataloguing trapping sites and calculating the point defect interactions for all individual boundaries across the entire misorientation range and for different axes. Indexing the possible segregation sites at the boundary and calculating realistic distributions for their associated trapping potentials is important for determining equilibrium interstitial occupancies at the interfaces [38]. The interstitial sites in a bulk BCC lattice are represented schematically in Fig. 7. These sites can be determined automatically for any given atomistic structure using a Delaunay triangulation. Due to the intrinsic lattice distortions induced by the geometric boundary between materials grains and its associated peculiar pattern of possible atomic relaxations, each grain boundary typically hosts a variety of non-equivalent interstitial sites, differing in coordination and volume from the reference bulk lattice values. The routine `hydrogenate.py` contains the `Hydrogenate` class for indexing the interstitial sites in a lattice, decorating a lattice with hydrogen, and computing the volume of interstitial sites. The routine `hydrogenate.py` contains the `Hydrogenate` class for indexing the interstitial sites in a lattice, decorating a lattice with interstitial hydrogens, and computing the volume of the interstitial sites. An alternative, and more complete, framework described in Ref. [39] which possesses extended functionality can also be used.

For each interstitial site of a given grain boundary it is possible to define an elastic dipole tensor (cf. Refs. [40–42]), a local quantity of particular significance because it allows modelling the coupling of a point defect to the strain fields of isolated dislocations, e.g., in the manner prescribed by Ref. [43]. The elastic dipole tensor elements G_{ij} are defined (cf. Ref. [41]) as follows:

$$\left.\frac{\partial\sigma_{ij}}{\partial n_d}\right|_{\epsilon} = G_{ij}, \tag{2}$$

where σ_{ij} is the volume averaged stress tensor, and n_d is the point defect (e.g., hydrogen) concentration. To compute these components we implemented in the `imeall.calc_elast_dipole` module the "defect force" scheme described in Ref. [42]. In this procedure, a point defect is introduced inside a unit cell of the host

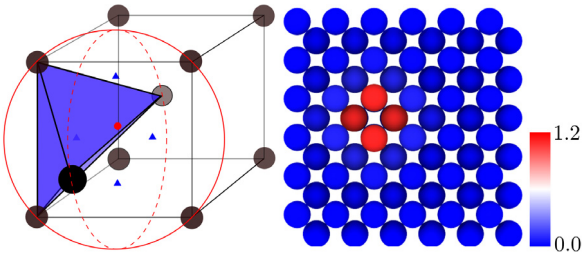

Fig. 7. Automatic determination of interstitial sites in a bulk BCC lattice using the Delaunay triangulation method. In a bulk BCC lattice the octahedral sites can be determined as the circumcentre of the sphere represented with the solid and dashed red lines. The solid red circle is centred in an octahedral interstitial site (solid red dot), and the blue triangular faces comprise a tetrahedral site, located at the centre of an irregular tetrahedron having vertices on four lattice sites. The right panel provides the magnitude of the forces induced by a hydrogen point defect in eV/Å units. These forces are required for calculation of the elastic dipole tensor.

material, and the structure is relaxed to its equilibrium zero-forces geometry. The defect is at this point removed from the model system keeping all other atoms fixed, and the resulting forces on the ions previously surrounding the interstitial are calculated without allowing any relaxation. These forces and the ion position vectors can then be related to the dipole tensor via the formula:

$$G_{ij} = \sum_{m \neq d} f_i^{[m,d]}(d_j - m_j), \tag{3}$$

where $\mathbf{d}$, $\mathbf{m}$ are the position vectors of the defect and host atoms, respectively, $\mathbf{f}^{[m,d]}$ is the force vector induced on the m-indexed atom by the removal of the d defect, and the subscripts i, j refer to spatial dimension components. We note that the defect force method does not require a single-Hamiltonian description (or an explicit total energy expression) for the system, so that forces can be computed using mixed ("embedding") schemes, i.e., combinations of descriptive potentials in the region of near crystallinity and more accurate quantum mechanical models at or very close to the point defect, for instance using the 'Learn On the Fly' scheme described in Ref. [44]. The absolute magnitude of the defect forces induced by a hydrogen interstitial in the BCC lattice is illustrated in Fig. 7, right panel, where as expected the majority of the calculated effect is limited to the metal ions neighbouring the interstitial site analysed.

The availability of the grain boundary database enables screening calculations on many defect geometries (e.g., distorted tetrahedral trapping sites located in the neighbourhood of the grain boundary) using fast classical potentials. It is therefore interesting to determine what accuracy level could be expected by the fast force models used in this screening step. We thus computed the dipole tensor for a tetrahedral site in bulk BCC Fe using an EAM force field [20] and a reference DFT calculation. The computed dipole tensor for the EAM is

$$\begin{pmatrix} 4.08 & 0 & 0 \\ 0 & 4.08 & 0 \\ 0 & 0 & 3.12 \end{pmatrix} \text{eV}, \tag{4}$$

and for the density functional calculation,

$$\begin{pmatrix} 4.65 & 0 & 0 \\ 0 & 4.65 & 0 \\ 0 & 0 & 3.52 \end{pmatrix} \text{eV}. \tag{5}$$

The DFT calculation was performed using the VASP package [45], using the PBE functional approximation to treat exchange and correlation [46], a PAW pseudopotential [47], a $3 \times 3 \times 3$ k-point mesh and imposing a 45 Ry energy cutoff on the plane wave wavefunction expansion. The periodic unit cell contained 250 Fe atoms with the H defect atom placed in a tetrahedral interstitial site in the middle of the unit cell. These results suggest a (∼12%) underestimation of the tensor elements size by the EAM potential. Such first order error could have a significant impact on a theoretical analysis of the interaction of the point defect with an elastic strain field present in the metal matrix; for instance in the vicinity of a grain boundary or a dislocation. However, such concern would be lifted by computing more precise values with higher accuracy DFT-based calculations only for a subset of most relevant/interesting cases revealed by an initial EAM-based high-throughput screening analysis, since systematic absolute ∼10% errors could be tolerated, counting on error cancellations, in the initial screening used to identify the subset. The ability to catalogue quantities such as the elastic dipole tensor according to the potential used, and to identify artefacts resulting from the model (e.g., by comparing different models and pointing out outliers) is in fact a useful additional function of the `Imeall` package. Similar calculations on the energetics of interstitials can be performed on any desired grain boundary structures and will be reported elsewhere.

7. Conclusion

We have described the structure and function of the `Imeall` package. The introduction of a naming convention and the overall structure of the database and specification of the data models provides a very convenient framework for a computational resource relating to interfacial structures in materials. The capabilities of the resource have been demonstrated with reference to various properties of symmetric tilt boundaries and pure crystalline α-Fe. The resource is offered online as fully open-access and is extensible by any user. The code repository can be found at https://github.com/kcl-tscm/imeall, links to the full structure database, which is hosted on the NOMAD servers and the web framework can be found in the documentation of the package at http://kcl-tscm.github.io/imeall/index.html. The database is stored on the NOMAD servers and can be navigated indirectly via http://imeall.co.uk/ or directly via https://labdev-nomad.esc.rzg.mpg.de/industry-project-imeall/. The server space is reserved for the next 10 years which provides a mid- to long term platform for storage of structures and public accessibility.

Acknowledgements

We would like to thank Prof. Adrian Sutton for illuminating conversations on grain boundaries and the physics of strain fields in their vicinity. We would also like to thank the anonymous referee for a suggestion to improve the naming convention. We would also like to thank Dr. Thomas Daff and Prof. G. Csányi for making their Fe GAP potential available for comparison. The GAP software is available for non-commercial use at www.libatoms.org. Financial support was provided by the Engineering and Physical Sciences Research Council under the HEmS program grant EP/L014742/1 and grant EP/P002188/1. This research used resources of the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC02-06CH11357. We are grateful to the UK Materials and Molecular Modelling Hub for computational resources, which is partially funded by EPSRC (EP/P020194/1). JRK, AF, and ADV acknowledge funding from the European Union's Horizon 2020 research and innovation program (Grant No. 676580, The NOMAD Laboratory, a European Centre of Excellence).

References

[1] A. Vattré, T. Jourdan, H. Ding, M.-C. Marinica, M.J. Demkowicz, Nature Commun. 7 (2016) 10424. URL http://dx.doi.org/10.1038/ncomms10424.

[2] X. Zhou, J. Song, Mater. Lett. 196 (2017) 123–127. http://dx.doi.org/10.1016/j.matlet.2017.03.035. URL http://www.sciencedirect.com/science/article/pii/S0167577X17303695.

[3] D. Bachurin, D. Weygand, P. Gumbsch, Acta Mater. 58 (16) (2010) 5232–5241. http://dx.doi.org/10.1016/j.actamat.2010.05.037. URL http://www.sciencedirect.com/science/article/pii/S1359645410003228.

[4] D.E. Spearot, M.D. Sangid, Slip Localization and Transfer in Deformation and Fatigue of Polycrystals, Curr. Opin. Solid State Mater. Sci. 18 (4) (2014) 188–195. http://dx.doi.org/10.1016/j.cossms.2014.04.001. URL http://www.sciencedirect.com/science/article/pii/S1359028614000175.

[5] P.D. Bristowe, A.G. Crocker, Phil. Mag. 31 (3) (1975) 503–517. http://dx.doi.org/10.1080/14786437508226533.

[6] V. Vitek, D.A. Smith, R.C. Pond, Phil. Mag. A 41 (5) (1980) 649–663. http://dx.doi.org/10.1080/01418618008239340.

[7] M. Tschopp, G. Tucker, D. McDowell, Acta Mater. 55 (11) (2007) 3959–3969. http://dx.doi.org/10.1016/j.actamat.2007.03.012. URL http://www.sciencedirect.com/science/article/pii/S1359645407001966.

[8] A.P. Sutton, R.W. Balluffi, Interfaces in Crystalline Materials, Oxford University Press, Great Clarendon Street, Oxford OX2 6DP, 1995.

[9] D.C. Handscomb, Canad. J. Math. 10 (1958) 85.

[10] P. Zeiner, Z. Kristallogr. 220 (2005) 915–925.

[11] R. Goldman, Synth. Lect. Comput. Graphics Animation 4 (1) (2010) 1–157. http://dx.doi.org/10.2200/S00292ED1V01Y201008CGR013.

[12] H. Grimmer, W. Bollmann, D.H. Warrington, Acta Crystallogr. Sect. A 30 (2) (1974) 197–207. http://dx.doi.org/10.1107/S056773947400043X.

[13] H. Grimmer, Acta Crystallogr. Sect. A 40 (2) (1984) 108–112. http://dx.doi.org/10.1107/S0108767384000246.

[14] J.D. Rittner, D.N. Seidman, Phys. Rev. B 54 (1996) 6999–7015. http://dx.doi.org/10.1103/PhysRevB.54.6999. URL http://link.aps.org/doi/10.1103/PhysRevB.54.6999.

[15] D.E. Spearot, M.A. Tschopp, K.I. Jacob, D.L. McDowell, Acta Mater. 55 (2) (2007) 705–714. http://dx.doi.org/10.1016/j.actamat.2006.08.060. URL http://www.sciencedirect.com/science/article/pii/S1359645406006537.

[16] M.A. Tschopp, D.L. Mcdowell, Phil. Mag. 87 (25) (2007) 3871–3892. http://dx.doi.org/10.1080/14786430701455321.

[17] M.A. Tschopp, K.N. Solanki, F. Gao, X. Sun, M.A. Khaleel, M.F. Horstemeyer, Phys. Rev. B 85 (2012) 064108. http://dx.doi.org/10.1103/PhysRevB.85.064108. URL http://link.aps.org/doi/10.1103/PhysRevB.85.064108.

[18] G.H. Campbell, S.M. Foiles, P. Gumbsch, M. Rühle, W.E. King, Phys. Rev. Lett. 70 (1993) 449–452. http://dx.doi.org/10.1103/PhysRevLett.70.449. URL https://link.aps.org/doi/10.1103/PhysRevLett.70.449.

[19] A.T. Paxton, J. Phys. D: Appl. Phys. 29 (7) (1996) 1689. URL http://stacks.iop.org/0022-3727/29/i=7/a=003.

[20] A. Ramasubramaniam, M. Itakura, E.A. Carter, Phys. Rev. B 79 (2009) 174101. http://dx.doi.org/10.1103/PhysRevB.79.174101. URL http://link.aps.org/doi/10.1103/PhysRevB.79.174101.

[21] E. Bitzek, P. Koskinen, F. Gähler, M. Moseler, P. Gumbsch, Phys. Rev. Lett. 97 (2006) 170201. http://dx.doi.org/10.1103/PhysRevLett.97.170201. URL http://link.aps.org/doi/10.1103/PhysRevLett.97.170201.

[22] M.I. Mendelev, S. Han, D.J. Srolovitz, G.J. Ackland, D.Y. Sun, M. Asta, Phil. Mag. 83 (35) (2003) 3977–3994 URL http://dx.doi.org/10.1080/14786430310001613264.

[23] S.L. Dudarev, P.M. Derlet, J. Phys.: Condens. Matter 17 (44) (2005) 7097. URL http://stacks.iop.org/0953-8984/17/i=44/a=003.

[24] H. Chamati, N. Papanicolaou, Y. Mishin, D. Papaconstantopoulos, Surf. Sci. 600 (9) (2006) 1793–1803. http://dx.doi.org/10.1016/j.susc.2006.02.010. URL http://www.sciencedirect.com/science/article/pii/S0039602806001701.

[25] A.P. Bartók, M.C. Payne, R. Kondor, G. Csányi, Phys. Rev. Lett. 104 (2010) 136403. http://dx.doi.org/10.1103/PhysRevLett.104.136403. URL https://link.aps.org/doi/10.1103/PhysRevLett.104.136403.

[26] A.P. Bartók, R. Kondor, G. Csányi, Phys. Rev. B 87 (2013) 184115. http://dx.doi.org/10.1103/PhysRevB.87.184115. URL https://link.aps.org/doi/10.1103/PhysRevB.87.184115.

[27] W.T. Read, W. Shockley, Phys. Rev. 78 (1950) 275–289. http://dx.doi.org/10.1103/PhysRev.78.275. URL https://link.aps.org/doi/10.1103/PhysRev.78.275.

[28] G.J. Ackland, A.P. Jones, Phys. Rev. B 73 (2006) 054104. http://dx.doi.org/10.1103/PhysRevB.73.054104. URL https://link.aps.org/doi/10.1103/PhysRevB.73.054104.

[29] D. Dragoni, T.D. Daff, G. Csányi, N. Marzari, Phys. Rev. Materials 2 (2018) 013808. http://dx.doi.org/10.1103/PhysRevMaterials.2.013808. URL https://link.aps.org/doi/10.1103/PhysRevMaterials.2.013808.

[30] A. Stukowski, K. Albe, Modelling Simulation Mater. Sci. Eng. 18 (8) (2010) 085001. URL http://stacks.iop.org/0965-0393/18/i=8/a=085001.

[31] A. Stukowski, V.V. Bulatov, A. Arsenlis, Modelling Simulation Mater. Sci. Eng. 20 (8) (2012) 085007. URL http://stacks.iop.org/0965-0393/20/i=8/a=085007.

[32] C. Hartley, Y. Mishin, Acta Mater. 53 (5) (2005) 1313–1321. http://dx.doi.org/10.1016/j.actamat.2004.11.027. URL http://www.sciencedirect.com/science/article/pii/S1359645404007062.

[33] J.P. Hirth, Metall. Trans. A 11 (6) (1980) 861–890. http://dx.doi.org/10.1007/BF02654700.

[34] R. Kirchheim, Acta Metall. 30 (6) (1982) 1069–1078. http://dx.doi.org/10.1016/0001-6160(82)90003-7. URL http://www.sciencedirect.com/science/article/pii/0001616082900037.

[35] R. Kirchheim, Prog. Mater. Sci. 32 (4) (1988) 261–325. http://dx.doi.org/10.1016/0079-6425(88)90010-2. URL http://www.sciencedirect.com/science/article/pii/0079642588900102.

[36] Y.A. Du, J. Rogal, R. Drautz, Phys. Rev. B 86 (2012) 174110. http://dx.doi.org/10.1103/PhysRevB.86.174110. URL http://link.aps.org/doi/10.1103/PhysRevB.86.174110.

[37] M. Yamaguchi, J. Kameda, K.-I. Ebihara, M. Itakura, H. Kaburaki, Phil. Mag. 92 (11) (2012) 1349–1368. http://dx.doi.org/10.1080/14786435.2011.645077.

[38] R.G.A. Veiga, M. Perez, C.S. Becquart, C. Domain, J. Phys.: Condens. Matter 25 (2) (2013) 025401. URL http://stacks.iop.org/0953-8984/25/i=2/a=025401.

[39] A. Goyal, P. Gorai, H. Peng, S. Lany, V. Stevanović, Comput. Mater. Sci. 130 (2017) 1–9. http://dx.doi.org/10.1016/j.commatsci.2016.12.040. URL http://www.sciencedirect.com/science/article/pii/S0927025617300010.

[40] M.J. Gillan, J. Phys. C: Solid State Phys. 17 (9) (1984) 1473. URL http://stacks.iop.org/0022-3719/17/i=9/a=006.

[41] D.A. Freedman, D. Roundy, T.A. Arias, Phys. Rev. B 80 (2009) 064108. http://dx.doi.org/10.1103/PhysRevB.80.064108. URL http://link.aps.org/doi/10.1103/PhysRevB.80.064108.

[42] R. Nazarov, J.S. Majevadia, M. Patel, M.R. Wenman, D.S. Balint, J. Neugebauer, A.P. Sutton, Phys. Rev. B 94 (2016) 241112. http://dx.doi.org/10.1103/PhysRevB.94.241112. URL https://link.aps.org/doi/10.1103/PhysRevB.94.241112.

[43] A. Cochardt, G. Schoek, H. Wiedersich, Acta Metall. 3 (6) (1955) 533–537. http://dx.doi.org/10.1016/0001-6160(55)90111-5. URL http://www.sciencedirect.com/science/article/pii/0001616055901115.

[44] G. Csányi, T. Albaret, M.C. Payne, A. De Vita, Phys. Rev. Lett. 93 (2004) 175503. http://dx.doi.org/10.1103/PhysRevLett.93.175503. URL https://link.aps.org/doi/10.1103/PhysRevLett.93.175503.

[45] G. Kresse, J. Furthmüller, Phys. Rev. B 54 (1996) 11169–11186. http://dx.doi.org/10.1103/PhysRevB.54.11169. URL https://link.aps.org/doi/10.1103/PhysRevB.54.11169.

[46] J.P. Perdew, K. Burke, M. Ernzerhof, Phys. Rev. Lett. 77 (1996) 3865–3868. http://dx.doi.org/10.1103/PhysRevLett.77.3865. URL https://link.aps.org/doi/10.1103/PhysRevLett.77.3865.

[47] P.E. Blöchl, Phys. Rev. B 50 (1994) 17953–17979. http://dx.doi.org/10.1103/PhysRevB.50.17953. URL https://link.aps.org/doi/10.1103/PhysRevB.50.17953.