Acta Materialia 125 (2017) 145–155

Contents lists available at ScienceDirect

Acta Materialia

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

Full length article

Efficient sampling in materials simulation - Exploring the parameter space of grain boundaries

Holger Dette [a], Josua Gösmann [a], Christian Greiff [a], Rebecca Janisch [b, *]

[a] Department of Mathematics, Institute of Statistics, Ruhr-Universität Bochum, 44780 Bochum, Germany
[b] Interdisciplinary Centre for Advanced Materials Simulation (ICAMS), Ruhr-Universität Bochum, 44780 Bochum, Germany

ARTICLE INFO

Article history:
Received 28 July 2016
Received in revised form
19 October 2016
Accepted 17 November 2016

Keywords:
Sequential sampling
Design of experiment
Grain boundary energy
Atomistic simulations
High-throughput methods

ABSTRACT

In the framework of materials design there is the demand for extensive databases of specific materials properties. In this work we suggest an improved strategy for creating future databases, especially for extrinsic properties that depend on several material parameters. As an example we choose the energy of grain boundaries as a function of their geometric degrees of freedom. The construction of many existing databases of grain boundary energies in face-centred and body centred cubic metals relied on the a-priori knowledge of the location of important cusps and maxima in the five-dimensional energy landscape, and on an as-densely-as-possible sampling strategy. We introduce two methods to improve the current state of the art. Based on an existing energy model the location and number of the energy minima along which the hierarchical sampling takes place is predicted from existing data points without any a-priori knowledge, using a predictor function. Furthermore we show that in many cases it is more efficient to use a sequential sampling in a “design of experiment” scheme, rather than sampling all observations homogeneously in one batch. This sequential design exhibits a smaller error than the simultaneous one, and thus can provide the same accuracy with fewer data points. The new strategy should be particularly beneficial in the exploration of grain boundary energies in new alloys and/or non-cubic structures.

1. Introduction

Increasing computational power contributes to the creation and extension of huge materials databases containing a variety of material properties. Such data collections can form the basis for material design and discovery, and thus promise great savings in development time and costs. The materials project [1] or the NOMAD repository [2] which together feature roughly $3.5 \cdot 6$ entries are impressive examples of such property collections. They can be used to search for correlations in property maps, or as input for material models.

High-throughput numerical simulations as well as experiments play a key role in determining materials properties, because they allow a systematic variation of individual parameters, such as e.g. composition, and the coverage of a broad range of these parameters. Nevertheless, the majority of properties that are available today are so-called intrinsic properties like phase stability, stiffness, or band gaps. Extrinsic properties which depend on additional structural parameters, like interface distribution functions, energies or mobilities, are available, too (see e.g. Ref. [3]), but much less frequently. As a matter of fact, because they depend on several variables at the same time, such properties form multidimensional databases of their own for one particular material.

For such cases the paper at hand shall promote an efficient sampling of the variable space that is based on design of experiment principles. As a suitable example we pick grain boundary energies, which depend on five geometric degrees of freedom.

The evolution of a microstructure of a metallic or multiphase material during solidification or deformation depends to a large extend on the energy of the interfaces in the microstructure. So do the resulting macroscopic properties, the deformability, thermal stability, and strength. Substantial scientific effort is put into describing the interactions of individual defects that form the microstructure within and across different material length scales to model their influence on the macroscopic properties. Again, numerical simulations are a valuable tool in this respect, because they allow the isolation of specific effects on different length or time

* Corresponding author.
E-mail address: rebecca.janisch@rub.de (R. Janisch).

http://dx.doi.org/10.1016/j.actamat.2016.11.044
1359-6454/

scales, as well as the abovementioned systematic variation of parameters that enter any material model.

In this paper we focus on grain boundaries, i.e. homophase interfaces, which separate two regions of undistorted crystal of the same crystal structure and chemistry that are misoriented with respect to each other. The space of macroscopic geometric parameters that define a grain boundary structure is five-dimensional - three degrees of freedom for the misorientation of the grains, two for the interface inclination. It is not per se clear which sampling strategy is the best for this five dimensional space. An intuitive approach might be to sample the parameter space homogeneously as finely as possible, so that afterwards an interpolation between the points is possible without knowing the underlying function that connects them. Alternatively, one can rely on a so-called structure model, which helps to identify the slowly and quickly varying regions in the energy function. Ideally, it already provides a closed form expression for the energy as a function of the structural parameters. This latter approach has been recently used by Bulatov et al. [4] to find a universal function for the crystallography class of face-centred cubic (fcc) metals. It reduces the amount of required data points, but uses a-priori knowledge of the topography of the multi-dimensional space of macroscopic parameters defining the grain boundary geometry. In particular it requires the specification of the number and the location of the cusps and maxima in the energy landscape. The 3D rotational subspaces that represent energy minima can indeed be predicted based on crystal symmetry. To a large extend, especially for fcc metals, the positions of the cusps in the one-dimensional subspaces of energy as a function of misorientation for twist and tilt grain boundaries can be predicted fairly well based on a random grain-boundary model [5], or a lattice-matching approach [6]. However, they also depend on the details of the atomistic structure and hence on the nature of the interatomic bonding, which makes a prediction more difficult already for bcc metals [6], and especially for systems with directional bonds, such as diamond or silicon, no straight forward relationship exists [7].

In this paper we show that the method proposed by Bulatov et al. [4] can be significantly improved in at least two directions. First we develop an estimate, which does not require any knowledge of the number and the location of cusps and maxima of the energy function, but determines these parameters from the data. Secondly, we address the problem of efficient sampling in such experiments. In contrast to previous work, which considers curve fits from all available data (within a subspace) we propose a sequential approach. A first part of the sample is used to obtain a prediction for the energy function. Subsequent sampling is performed sequentially, such that the new experiment is conducted at sampling points at which the uncertainty of the prediction is maximal. Additionally, the prediction is continuously updated using the new information of the previous experiments. In many cases we can thus minimize the number of points which are necessary to give a reliable description of the parameter space, especially cusps and maxima in the energy landscape. Thus, the important parameters of the energy function do not have to be known a priori, and still the computational effort remains manageable.

The remaining part of this paper is organized as follows. In Section 2 we give a brief description of the geometry and structure of the grain boundary and describe the existing data bases of grain boundary energies. We also describe the structural model introduced in Ref. [4] which will be further developed in this paper. Section 3 introduces the new statistical methodology, and the advantages of the sequential sampling approach are illustrated in Section 4.

2. Grain boundary structure and energy

The geometry and structure of grain boundaries are determined by the five macroscopic and three microscopic degrees of freedom (DOF), i.e. by.

- The misorientation of the two grains, e.g. defined by rotation axis $\widehat{\boldsymbol{\rho}}$ and angle x, or a rotation matrix $\mathbf{A}$, which rotates grain 1 into grain 2. This rotation determines three macroscopic DOF.
- The inclination of the grain boundary plane, defined by its normal vector $\widehat{\mathbf{n}}$, that determines the remaining two macroscopic DOF.
- A translation $\mathbf{t}$ of the grains parallel and perpendicular to the interface plane, which fixes the three microscopic DOF.

Traditionally, grain boundaries for which the plane normal vector $\widehat{\mathbf{n}}$ is perpendicular to the misorientation axis $\widehat{\boldsymbol{\rho}}$ are called tilt grain boundaries, those with $\widehat{\mathbf{n}}$ parallel to $\widehat{\boldsymbol{\rho}}$ twist grain boundaries.

Several other representations of the grain boundary geometry are possible, e.g. a representation by two rotation matrices that describe the orientation of each grain with respect to a reference frame, in which one axis is perpendicular to the interface [8]. It depends on the application which representation is the most convenient one. For instance, a representation by two rotation matrices is useful to construct grain boundary structures as input for atomistic simulations, as done e.g. in Ref. [9]. However, to evaluate afterwards the energies of grain boundaries as a function of their degrees of freedom, as done e.g. by Wolf [10] for face centred cubic metals, a representation in terms of rotation axis and angle is more descriptive.

Some grain misorientations lead to a periodic superstructure across the grain boundary, a so-called “coincidence site lattice” (CSL). CSL-based grain boundaries represent low-energy structures that are especially important when trying to capture the material properties. Hence, besides being characterised by their misorientation they are usually described by an additional parameter Σ, which is the ratio of the CSL unit cell volume to that of the original crystal lattice unit cell. In other words it is a measure for the periodicity of the CSL lattice.

In principle the energies of all possible grain boundary structures in a microstructure could be determined from atomistic simulations. The only limiting factors are the computational demand and the availability of a reliable description of the interatomic interactions. Many atomistic studies of lower-dimensional subspaces can be found in the literature, which focus e.g. on symmetrical tilt grain boundaries (e.g. Refs. [10–13]), on the two-dimensional subspace of grain boundary plane inclinations [14], or on examples of general [15], and heterophase [16] boundaries. In addition, few but expansive databases which represent the full five-dimensional space have been created that include several thousand structures [9,17,18].

Up to date there exist two major collections of grain boundary energies which were obtained from atomistic simulations, the database of Kim et al. [17] for body-centred cubic (bcc) iron, and the one of Olmsted and co-workers [9] for face-centred cubic (fcc) aluminium and nickel, and a more recent version also for bcc iron [18]. The structure of the databases reflects two different approaches: Kim et al. sample the 5D parameter space as densely as possible in a homogeneous fashion, i.e. for arbitrary misorientations, to provide sufficient data for a numerical interpolation scheme. Details are given in Section 2.1. In the approach chosen in Ref. [9] the emphasis is put on the CSL-based grain boundaries by sampling the subspace of special misorientations around high-symmetry axes in the cubic lattice. Thus, in the resulting database