Inference and Uncertainty Propagation of GB Structure-Property Models: H Diffusivity in [100] tilt GBs in Ni

Oliver K. Johnson[a,*], Eric R. Homer[a], David T. Fullwood[a], David E. Page[a], Kathryn F. Varela[a], Sterling G. Baird[a]

[a]*Department of Mechanical Engineering, Brigham Young University, Provo, UT 84602, USA*

Abstract

In this work we present a non-parametric Bayesian approach for developing structure-property models for grain boundaries (GBs) with built-in uncertainty quantification (UQ). Using this method we infer a structure-property model for H diffusivity in [100] tilt GBs in Ni at 700 K based on molecular dynamics (MD) data. Once a GB structure-property model is developed, it can be used as an input to mesoscale simulations of the effective properties of polycrystals, microstructure evolution, etc. A significant advantage of the Bayesian approach presented here is that it facilitates propagation of uncertainties from the underlying structure-property model to the output predictions from mesoscale modeling. Leveraging this capability, we perform mesoscale simulations of the effective diffusivity of polycrystals to investigate the interaction between structure-property model uncertainties and GB network structure. We observe a fundamental interaction between crystallographic correlations and spatial correlations in GB networks that causes certain types of microstructures (those with large populations of J_2- and J_3-type triple junctions) to exhibit intrinsically larger uncertainty in their effective properties.

Keywords: Grain Boundary, Structure-Property Model, Bayesian Inference, Uncertainty Quantification, Uncertainty Propagation

1. Introduction

In attempting to develop structure-property models for grain boundaries (GBs), one is faced with several challenges, including: (i) the amount of available data is often small compared to the size of the GB character space; (ii) the existing data has finite accuracy and precision (i.e. it contains some amount of uncertainty); and (iii) the functional form of the structure-property model is generally unknown.

Data insufficiency is a perennial challenge for GBs because of the size of the 5D GB character space and the relatively high cost of GB characterization and property measurement (whether experimentally or computationally). The most commonly employed databases contain on the order of 10^2 GBs [1, 2]. In contrast, Rohrer et al. predicted that for cubic materials the number of GBs in a 5° grid over the GB character space would be on the order of 10^5 [3].

Although less frequently discussed than data quantity, the accuracy and precision of GB data is also of critical importance. All observations of GB properties, whether from calculations or from experimental observations, contain some degree of both epistemic and aleatoric uncertainties. The development of robust and reliable structure-property models requires that such uncertainties be incorporated and inform model predictions and their interpretation.

With regard to the functional form of GB structure-property models, this is still largely an

*Corresponding author.
Email address: ojohnson@byu.edu (O.K. Johnson).

Table 1: Summary of datasets employed in this work. All datasets are for Ni. Simulation methods include molecular statics (MS) and molecular dynamics (MD). For DS01 each of the MD simulations yielded one diffusion coefficient for the bulk and one for the respective GB.

Dataset	Property	GBs	Method	Reference
DS01	H Diffusivity	26 [100] symmetric tilts 26 bulk	MD	[4]
DS02	GB Energy	40 [100] asymmetric and symmetric tilts	MS	[5, 6]
DS03	GB Energy	4 [100] symmetric tilts	MS	[1]

open question. Read and Shockley developed a classical model to predict GB energy for low-angle GBs (LAGBs) [7]. Recently, an interpolation function based on the functional form of the Read-Shockley model was developed to predict GB energy for the full 5D GB character space for FCC metals [8]. However, for the vast majority of GB properties, physical theories have not yet been developed to predict the functional form of the structure-property model. Consequently, even when GB property data is available, in most cases a corresponding physically motivated parametric model whose coefficients can be fit to the data does not exist.

Bayesian inference techniques are an attractive option to address these challenges because they (i) can handle underdetermined systems; (ii) naturally incorporate data and modelization uncertainties; and (iii) can be formulated in a non-parametric way.

In the geostatistics community, Bayesian inference has been applied heavily. For example, in seismic tomography, models for the sub-surface geological structure of the earth (e.g. spatially resolved material composition) are reconstructed from travel times of seismic waves [9–12]. There are many parallels between such geophysical problems and the present problem of inferring GB structure-property models, including: measurements are often severely limited in number (due to cost and infrastructure) and contain uncertainty from a variety of sources, and an analytical ansatz for the model form may not be available. Consequently, probabilistic approaches have been developed that can handle limited data (even severely underdetermined systems) [13, 14], and can infer continuous models from discrete data [13, 14], in non-parametric ways (i.e. they do not require any ansatz of the model form) [14, 15]. Gaussian process regression methods [16–20], currently popular in machine learning applications, are a special case of these more general methods.

In this work, we adapt Bayesian techniques from the geostatistics community to provide a method to infer GB structure-property models and apply this approach to obtain a structure-property model for H diffusivity in [100] tilt GBs in Ni at 700 K with quantified uncertainty. We then use this model to perform uncertainty propagation in mesoscale simulations of GB network diffusivity in polycrystals to investigate the interaction between structure-property model uncertainty and GB network structure.

2. Methods

2.1. Datasets

For our structure-property model inference, we combine several datasets including our own calculations as well as other datasets published in the literature (see Table 1). The first dataset (DS01) consists of both bulk and GB hydrogen diffusivity values for 26 [100] symmetric tilt GBs in Ni, calculated via molecular dynamics (MD) [4].

In the same reference [4], we recently demonstrated that hydrogen diffusivity in Ni GBs can be predicted from the corresponding 0 K GB energy via an extension of the Borisov relation [21–23], with appropriately calibrated parameters. This enables us to expand our pool of data to leverage GB energy data, which is much more ubiquitous than direct diffusivity values. Therefore, we also employ several datasets containing GB energy values

for [100] tilt GBs, including our own [5, 6] (DS02), and other extant datasets from the literature [1] (DS03). In Appendix A we describe our process for predicting GB diffusivity from GB energy with quantified uncertainty.

2.2. Data Uncertainty

The Bayesian inference approach we employ requires the specification of measurement/calculation uncertainty for all of the input data.

2.2.1. Diffusivity

For the direct MD diffusivity data (DS01), the diffusivity is calculated from the mean-squared-displacement (MSD), $\langle R^2 \rangle$, of the hydrogen atoms over time according to (see, e.g., [24])

$$D = \frac{1}{6} \frac{d \langle R^2 \rangle}{dt} \tag{1}$$

In this case, the aleatoric (stochastic) uncertainty can be quantified via a bootstrapping method in which the total simulation time is divided into disjoint sub-intervals and the diffusivity is determined in each of these sub-intervals [25]. Figure 1 illustrates this process and shows the distribution of diffusivity values obtained in this manner for a 12.7° [100] symmetric tilt GB. We use 5 disjoint sub-intervals over the last half of the simulation time so as to avoid the standard initial non-linear transient. Figure 2 shows the distribution of diffusivity values thus obtained for all of the 26 [100] symmetric tilt GBs in dataset DS01.

The epistemic (modeling) uncertainty for MSD calculations is due primarily to the interatomic potential employed. Comparisons of calculated values to experiments reveal that MSD diffusivities for H in FCC metals are often about 2 times higher than experimentally measured values [4, 26]. At the same time, typical aleatoric uncertainties in experimental measurements are themselves on the order of 5 % – 65 % [27–30], and for the listed references the mean uncertainty is 25 %. In contrast the aleatoric uncertainties in our own MSD calculations of GB diffusivity are in the range 3 % – 15 %, with the mean uncertainty being 7.5 %. While the MSD results are more precise, it is not obvious whether the experimentally measured or calculated values

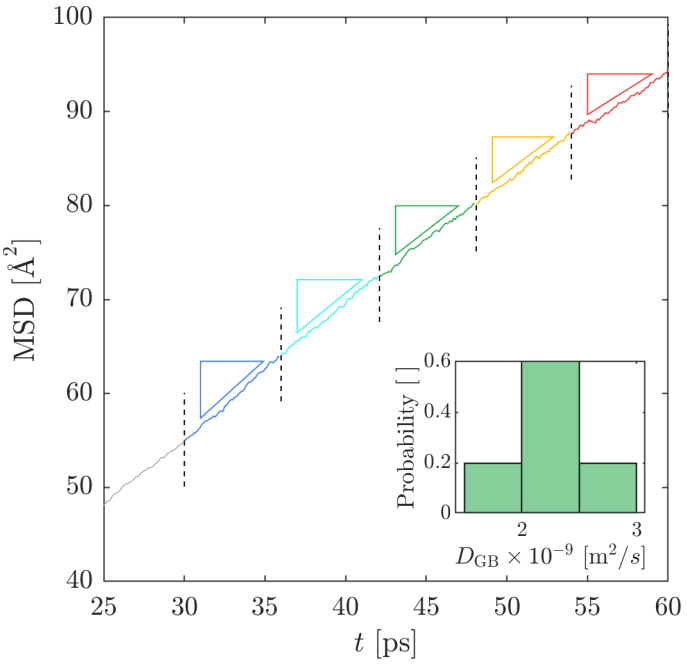

Figure 1: Illustration of the bootstrapping process to obtain estimates of bulk and GB diffusion coefficients and their corresponding uncertainty from MD simulations. The slope of MSD vs. t in each sub-interval provides an estimate of the diffusion coefficient. The uncertainty is calculated from the resulting distribution (see inset).

are more accurate. Moreover, the epistemic uncertainty arising from the discrepancy between calculated and measured values is generally not well represented by a distribution [25], and has not yet been sufficiently well quantified to be considered. Consequently, we will neglect this epistemic uncertainty. If one desires, the final results can be scaled to compensate for the discrepancy between calculated and experimental values, depending on which of the two the user believes to be more accurate.

2.2.2. Energy

For the MD GB energy data, sources of error include the interatomic potential, and the minimization procedure. Unfortunately, uncertainty in GB energy is rarely reported. For symmetric tilt GBs in UO_2, Ksibi et. al report GB energy uncertainties of 2.4 % – 4 % at 1700 K [31]. For Ni at 0 K (the present system), explicit uncertainties are not available; however, several authors suggest that errors on the order of 5 % – 10 % are reasonable [8, 32]. Consequently, we assume a conservative uncertainty of 10 % for all MD GB energy

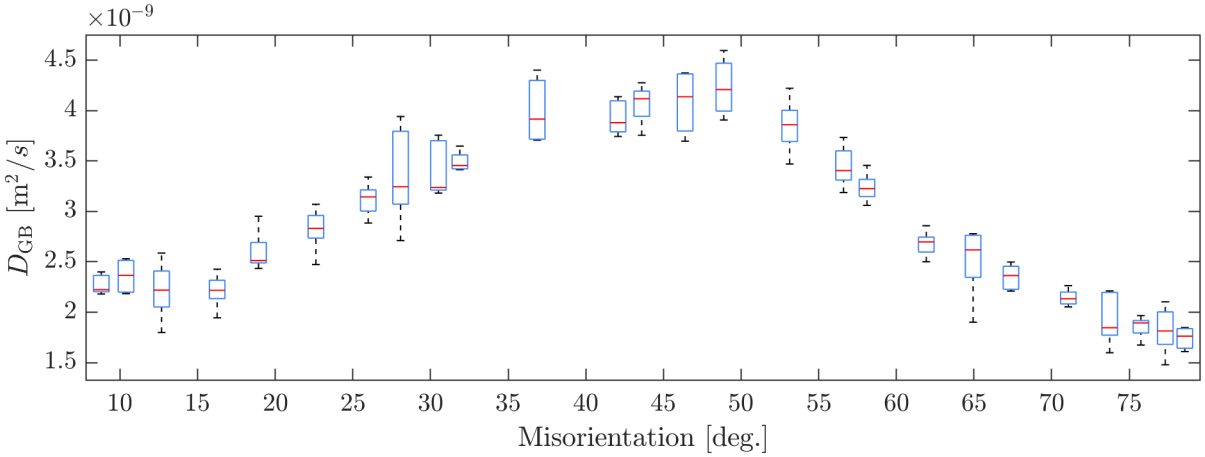

Figure 2: Diffusion coefficients for the 26 [100] symmetric tilt GBs in DS01, together with quantified uncertainty.

values. Rohrer, et al. compared MD GB energies to experimentally measured GB energies and found good agreement for those GBs that were well represented in both datasets [32]. They also suggest that because the MD GB energies do not depend on observation frequency (unlike the experimental values in their comparison), they may be more reliable than the experimental values. Nevertheless, there are certainly important outstanding questions about, e.g. the impact of metastable GB states at finite temperatures in experimental samples and other differences between the conditions simulated in MD and those present in experimental measurements that likely impact the epistemic uncertainty of both types of data. Again, because suitable estimates are not yet available, we ignore any epistemic uncertainty of the MD data.

2.3. Modeling Uncertainties

One important additional point must be made regarding uncertainties. For both diffusivity and energy data, the uncertainty is almost universally reported in the form $\pm f$ where f is either an absolute error or a percent error. This representation of the uncertainty implicitly assumes that the errors are Gaussian (or at least come from a symmetric distribution). This is, however, inconsistent with the nature of the properties being measured since both GB diffusivity and energy are non-negative and a Gaussian distribution assigns finite probability to negative values. In reality the uncertainties in these properties must come from an asymmetric distribution whose support is $\mathbb{R} \geq 0$, such as the log-normal distribution[1]. However, when the dispersion is small the log-normal distribution approaches a normal distribution and we have that $P \sim \text{Lognormal}(\mu, \sigma^2) \approx \mathcal{N}(m, s^2)$, where P is the measured property of interest (energy, γ, or diffusivity, D), and the notation $a \sim B$ denotes that the random variable a follows a B distribution.

If m is the reported mean property value, and $s = fm$ is the reported standard deviation, where f is the percent error (i.e. $m \pm (f \cdot 100)\%$), the parameters of the corresponding log-normal distribution are given by

$$\mu = \ln\left(\frac{m}{\sqrt{f^2+1}}\right) \tag{2}$$

$$\sigma = \sqrt{\ln(f^2+1)} \tag{3}$$

The above distinction is not merely pedantic. If inference is performed, incorrectly assuming that the uncertainties in diffusivity or energy are Gaussian, models sampled from the posterior will yield non-physical negative values. To enforce the physical constraint of non-negativity for these properties, log-normal priors will be used[2]. At the same time,

[1] Technically the support of the log-normal distribution is $\mathbb{R} > 0$, but the range of diffusivity values typically observed makes this point practically irrelevant.

[2] Technically any distribution whose support is in $\mathbb{R} \geq 0$ may be used, but, as will be explained later, the choice of a log-normal distribution—if it is appropriate—is particularly convenient.

Gaussian distributions furnish a number of convenient closed-form results that make the inference process highly efficient. To satisfy the physical constraints of non-negativity and also take advantage of the convenient machinery of Gaussian distributions, the entire problem can be log-transformed by making use of the fact that if $P \sim \text{Lognormal}(\mu, \sigma^2)$ then $\ln P \sim \mathcal{N}(\mu, \sigma^2)$. Rather than performing inference on D and γ directly, we therefore perform the inference on $\ln D$ and $\ln \gamma$, and we convert reported uncertainty values in the form $\pm f$ to the desired form by means of Eq. 3.

2.4. Inference

Our inference approach is based on the Bayesian methods of Tarantola, et al. [11–15, 33, 34]. A detailed derivation is provided in our recent work [35]. We provide a brief summary here.

We will use bold font to abstractly denote any non-scalar objects including functions, operators, vectors, matrices, and sets. For example, $\mathbf{m}$ denotes the function $m(\omega, \beta)$, which may be implemented numerically as the vector defined by $m_i = m(\omega_i, \beta_i)$; and $\mathbf{C}$ denotes the operator $C(\mathbf{x}, \mathbf{x}')$, which may be implemented numerically as the matrix defined by $C_{ij} = C(\mathbf{x}_i, \mathbf{x}'_j)$.

Let $\mathscr{D} = \mathscr{D}(\omega, \beta)$ be the unknown structure-property model describing the diffusivity of [100] tilt GBs, where ω and β are the independent variables describing the GB misorientation and plane inclination, to be defined more explicitly later. We wish to infer $\mathscr{D}$ from observations of the diffusivity, $\mathbf{D}^{\text{obs}} = \{D_1^{\text{obs}}, D_2^{\text{obs}}, \ldots, D_N^{\text{obs}}\}$, and crystallography, $\mathbf{M} = \{M_1, M_2, \ldots, M_N\}$, of some number, N, of GBs, where D_n^{obs} is the diffusivity of the n-th GB characterized by crystallographic parameters $M_n \equiv \{\omega_n, \beta_n\}$. As $\mathscr{D} \in [0, \infty)$ we make the following transformations

$$\mathbf{m} = \ln(\mathscr{D}) \tag{4}$$

$$\mathbf{d}_0 = \ln(\mathbf{D}^{\text{obs}}) \tag{5}$$

$$g(\mathbf{m}) = \mathbf{m} \tag{6}$$

so that $\mathbf{m}, \mathbf{d}_0, g(\mathbf{m}) \in (-\infty, \infty)$. We assume Gaussian priors for both the data and the model, with mean, $\mathbf{d}_0$ and $\mathbf{m}_0$, and covariance, $\mathbf{C}_{\mathbf{d}_0}$ and $\mathbf{C}_{\mathbf{m}_0}$, respectively. Because the priors are Gaussian and $g(\mathbf{m})$ is a linear function of $\mathbf{m}$, the posterior will also be Gaussian and takes the particularly simple form [14]

$$\sigma_m(\mathbf{m}) = k \exp\left(-\frac{1}{2} S(\mathbf{m})\right) \tag{7}$$

where

$$\begin{aligned} S(\mathbf{m}) = {} & (g(\mathbf{m}) - \mathbf{d}_0)^\intercal \mathbf{C}_{\mathbf{d}_0}^{-1} (g(\mathbf{m}) - \mathbf{d}_0) \\ & + (\mathbf{m} - \mathbf{m}_0)^\intercal \mathbf{C}_{\mathbf{m}_0}^{-1} (\mathbf{m} - \mathbf{m}_0) \end{aligned} \tag{8}$$

is referred to as the misfit function. In general, the vector-valued[3] functional $g(\mathbf{m})$ represents the map between the model space, $\mathbb{m}$, and the data space, $\mathbb{d}$. If observations correspond to the effective properties of polycrystals as in [35, 36], then $g(\mathbf{m})$ is a homogenization equation. In the present case where only direct observations of individual GBs are considered we have the trivial expression given in Eq. 6.

$\sigma_m(\mathbf{m})$ is a probability density function that describes everything that is known about the model $\mathbf{m}$ from both our observations and priors, and therefore in the most general sense $\sigma_m(\mathbf{m})$ itself represents the solution to the inference problem. If, rather than the probability density over possible models, a particular model realization is desired, a natural (though not the only) choice is the model that maximizes the posterior probability density. Due to the Gaussian form of the posterior in the present case, the maximum *a posteriori* (MAP) model estimate coincides with the mean of the posterior, which occurs at [13–15]

$$\begin{aligned} \widetilde{\mathbf{m}} = {} & \mathbf{m}_0 + \mathbf{C}_{\mathbf{m}_0} \mathbf{G}^* (\mathbf{C}_{\mathbf{d}_0} + \mathbf{G} \mathbf{C}_{\mathbf{m}_0} \mathbf{G}^*)^{-1} \\ & \cdot (\mathbf{d}_0 - g(\widetilde{\mathbf{m}}) + \mathbf{G} (\widetilde{\mathbf{m}} - \mathbf{m}_0)) \end{aligned} \tag{9}$$

and is also the location of the minimum of $S(\mathbf{m})$. The covariance operator of the posterior is necessary to quantify the posterior uncertainty and can

[3] To follow the conventions used in the literature and avoid burdensome notation, $g(\mathbf{m})$ implicitly represents $g(\mathbf{m}, \mathbf{M})$—i.e. it is a vector in which the n-th element corresponds to the functional $g(\mathbf{m})$ being evaluated for the n-th microstructure, $\mathbf{M}_n$, making the arithmetic operation $g(\mathbf{m}) - \mathbf{d}_0$ appearing in Eq. 8 sensible.

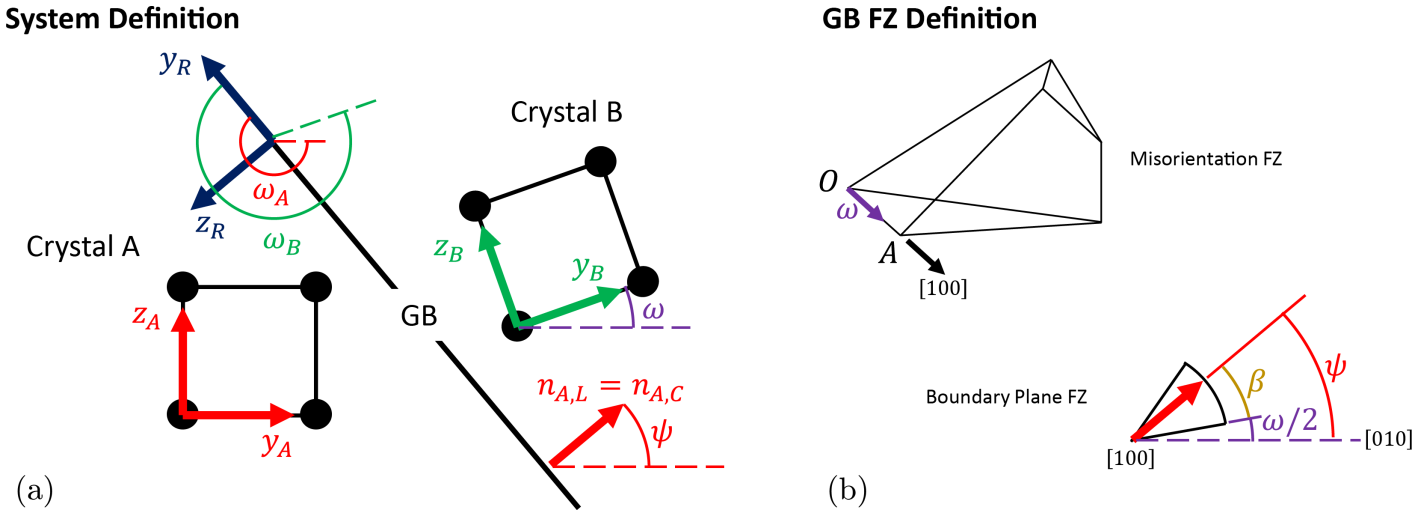

Figure 3: Visual definition of the crystallographic variables used in this work. In (a) a schematic illustration of an arbitrary [100] tilt GB is used to define the misorientation angle, ω, the polar angle of the boundary plane in the reference frame of crystal A, ψ, and the lattice orientations of grains A and B in the GB reference frame, ω_A and ω_B, respectively. In (b) the complete GB misorientation FZ for cubic crystal symmetry is shown with the subdomain of misorientations with [100] rotation axes highlighted. The boundary plane FZ is also shown and the polar angle of the GB plane, β is defined.

be calculated via [13–15]

$$\mathbf{C}_{\widetilde{\mathbf{m}}} = \left(\mathbf{I} - \mathbf{C}_{\mathbf{m}_0}\mathbf{G}^*(\mathbf{C}_{\mathbf{d}_0} + \mathbf{G}\mathbf{C}_{\mathbf{m}_0}\mathbf{G}^*)^{-1}\mathbf{G}\right)\mathbf{C}_{\mathbf{m}_0} \tag{10}$$

where $(\cdot)^*$ denotes the adjoint of $(\cdot)$ and $\mathbf{G}$ is the Fréchet derivative operator, evaluated at $\widetilde{\mathbf{m}}$. In general, the Fréchet derivative evaluated at a point $\mathbf{m}$ is defined by [14]

$$\mathbf{G}\delta\mathbf{m} = g(\mathbf{m} + \delta\mathbf{m}) - g(\mathbf{m}) \tag{11}$$

where $\delta\mathbf{m}$ is a small perturbation.

Equation 9 is an implicit equation (note that $\widetilde{\mathbf{m}}$ appears on both the right- and left-hand sides, and that $\mathbf{G}$ depends on $\widetilde{\mathbf{m}}$), and can be solved by an iterative fixed-point method [15]

$$\widetilde{\mathbf{m}}_{k+1} = \mathbf{m}_0 + \mathbf{C}_{\mathbf{m}_0}\mathbf{G}_k^*(\mathbf{C}_{\mathbf{d}_0} + \mathbf{G}_k\mathbf{C}_{\mathbf{m}_0}\mathbf{G}_k^*)^{-1} \cdot (\mathbf{d}_0 - g(\widetilde{\mathbf{m}}_k) + \mathbf{G}_k(\widetilde{\mathbf{m}}_k - \mathbf{m}_0)) \tag{12}$$

where $\mathbf{G}_k$ denotes the the Fréchet derivative evaluated at $\widetilde{\mathbf{m}}_k$. We choose the prior model $\mathbf{m}_0$ as the starting point (though the solution is independent of this choice), and convergence typically occurs within only a few iterations (for the results presented here only 11 iterations were required).

2.5. Defining the Fundamental Zone

For both computational efficiency, and proper design of the prior model covariance function (discussed below), it is necessary to explicitly define the fundamental zone (FZ) over the relevant domain.

Consider the [100] tilt GB shown in Fig. 3(a), coordinated by two crystals with their respective coordinate systems shown. Without loss of generality, we assume that crystal A is aligned with the laboratory reference frame so that the GB normal expressed in the lab frame is equivalent to the GB normal expressed in the crystal frame of grain A as shown[4]. The misorientation angle is denoted by ω and the polar angle of the GB normal in the reference frame of crystal A is denoted by ψ.

Figure 3(b) shows the complete misorientation FZ for cubic symmetry. The sub-manifold corresponding to misorientations having a [100] rotation axis is the line OA where ω increases from 0° to 45°. The corresponding boundary plane FZ is a stereographic triangle rotated by an angle $\omega/2$ from the [010] axis (see [37, 38]). The subset of [100] tilts inhabit the outer arc of this stereographic triangle and may be parameterized by the polar angle

[4]Note that we employ the active rotation convention for all calculations

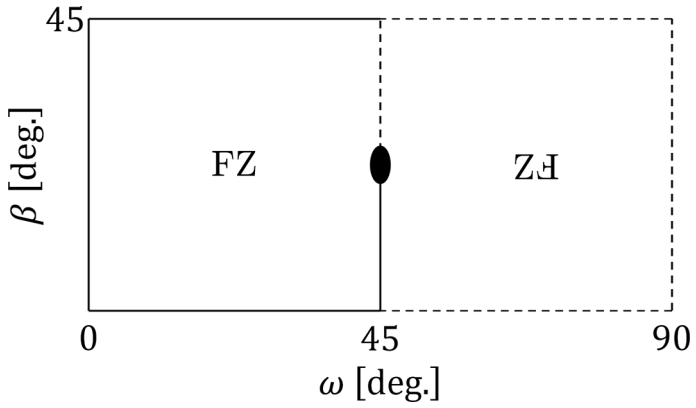

Figure 4: The fundamental zone (FZ) for [100] tilt GBs for cubic crystal symmetry. The left portion with $\omega \in [0°, 45°]$ and $\beta \in [0°, 45°]$ is the FZ. Note that the FZ is not closed as the line $\omega = 45°$ over the range $\beta \in [22.5°, 45°]$ is *not* included in the FZ. The oval at the point $(\omega, \beta) = (45°, 22.5°)$ represents a two-fold axis of symmetry. When the domain $\omega \in [0°, 90°]$ and $\beta \in [0°, 45°]$ is considered it is referred to as FZ2 since it contains two symmetric copies of the FZ.

$\beta \in [0°, 45°]$. As indicated in the figure, this angle is related to the polar angle of the GB normal expressed in the reference frame of crystal A via $\beta = \psi - \omega/2$. Combining the misorientation and boundary plane FZ subspaces we arrive at the complete FZ for [100] tilt GBs defined by $\omega \in [0°, 45°]$ and $\beta \in [0°, 45°]$, which is shown in Fig. 4. Mirror planes exist for $\omega = 0°$, $\omega = 90°$, $\beta = 0°$, and $\beta = 45°$, and there is a two-fold rotation symmetry about the point $(\omega, \beta) = (45°, 22.5°)$. In consequence of this two-fold rotation, this FZ is often displayed together with a neighboring symmetric copy as shown in Fig. 4 (see e.g. [8]). We will refer to this augmented domain defined by $\omega \in [0°, 90°]$ and $\beta \in [0°, 45°]$ as FZ2. One common implicit use of FZ2 in the literature is for the display of the subset of symmetric [100] tilts, which can be compactly represented by the line $\beta = 0$ with $\omega \in [0°, 90°]$, as we have done in Fig. 2.

2.6. Designing the Covariance (Kernel) Function

Structure-property models for GBs must respect certain physical invariances (e.g. crystallographic symmetries). These constraints represent important prior information about possible models that must inform the design of proper priors. There are two types of invariance that are relevant to the present problem.

2.6.1. Crystallographic Symmetry

The first type of invariance consists of crystallographic symmetry operations. With the domain of interest restricted to FZ2, there remains a finite group, H, of crystallographic symmetry operations consisting of the identity and a two-fold rotation about the point $(\omega, \beta) = (45°, 22.5°)$. Ginsbourger et al. [39, 40] showed that a random field can be invariant under a group action if and only if the corresponding covariance function (commonly referred to as a *kernel*) is argument-wise invariant under that group action. This means that for models resulting from our posterior distribution—including the MAP estimate or any random samples from the posterior distribution—to exhibit the proper symmetries, we must design a covariance function that exhibits those symmetries.

One way to construct such a covariance function that will be familiar to the microstructure community involves the use of a properly symmetrized distance function. As we will show below, this approach succeeds in enforcing the desired symmetries, but at the cost of introducing topological artifacts. We propose an alternative approach that circumvents this drawback by employing an unsymmetrized distance function in conjunction with the sum-over-group-orbits strategy [39–41].

Distance metrics for GBs have been an active area of research for some time and there are a variety of candidates in the literature (see e.g. [42–46] and [47] for a recent review of extant distance metrics). We choose to formulate our approach in the context of the recently developed octonion distance metric [45].

A GB can be represented by a unit octonion defined by [45]

$$o = \frac{1}{\sqrt{2}} [q_A, q_B] \tag{13}$$

where q_A and q_B are unit quaternions representing the crystal orientations of grains A and B incident to the GB, expressed in the GB plane reference

frame[5] [45], and defined by

$$q = \left[\cos\left(\frac{\omega}{2}\right), \sin\left(\frac{\omega}{2}\right)\hat{\mathbf{n}}\right] \tag{14}$$

where ω and $\hat{\mathbf{n}}$ are the angle and axis of rotation, respectively. The distance between two GB octonions (GBOs) is computed via [45]

$$\begin{aligned}\Omega &= 2\arccos|o_1 \cdot o_2| \\ &= 2\arccos\left(\frac{1}{2}|q_A \cdot q_C + q_B \cdot q_D|\right)\end{aligned} \tag{15}$$

where the first GB is coordinated by grains A and B and the second GB is coordinated by grains C and D. Due to the existence of a free parameter, ζ, in this representation (rigid body rotations of the GB about its normal leave it unchanged) one must normally minimize this distance with respect to ζ. However, for [100] tilt GBs $\zeta_{\min} = 0$ and no minimization is necessary. When crystal symmetries exist, all symmetrically equivalent GBOs are considered and the smallest distance is chosen.

For [100] tilt GBs Eq. 15 simplifies to

$$\Omega = 2\arccos\left|\cos\left(\frac{\omega_{AB} - \omega_{CD}}{4}\right)\cos\left(\frac{\beta_A - \beta_C}{2}\right)\right| \tag{16}$$

where ω_{XY} is the misorientation angle between grains X and Y and β_X is the GB inclination parameter for the GB normal pointing away from grain X (outward-pointing normal convention). We note that for symmetric [100] tilts we have $\beta_A = \beta_C = 0$ and the GBO distance is equal to half the difference in tilt angles.

With a distance function defined we introduce the resulting covariance (kernel) function. There are numerous choices available depending on the application and one's assumptions about the nature of the function to be inferred (e.g. differentiability). A common choice is the Gaussian (or squared-exponential) covariance

$$C(\mathbf{x}, \mathbf{x}') = \sigma^2 \exp\left(-\frac{d^2(\mathbf{x}, \mathbf{x}')}{2L^2}\right) \tag{17}$$

where σ is a constant marginal uncertainty (standard deviation), L is the "smoothness length," and $d^2(\mathbf{x}, \mathbf{x}')$ is the squared distance between points $\mathbf{x}$ and $\mathbf{x}'$; for the present problem $\mathbf{x} \equiv (\omega, \beta)$. The Gaussian covariance is stationary and, when the distance function is Euclidean (in any dimension), this covariance function is positive definite [48, 49]. However, in many cases (like the present) the physically relevant distance between two points is non-Euclidean, and in such cases the guarantee of positive-definiteness is lost and $C(\mathbf{x}, \mathbf{x}')$ ceases to be a valid covariance function.

There are two common strategies for addressing this: (i) multidimensional scaling, and (ii) approximation. In the multidimensional scaling approach, one uses dimensionality reduction techniques in reverse to find a mapping to a sufficiently *high*-dimensional space that the distance becomes Euclidean [50]. Alternatively, one can find the closest symmetric positive-definite covariance function under some appropriate norm, e.g. Frobenius [51]. We adopt the latter approach, which can be readily computed in MATLAB via the `nearestSPD()` function [52].

There are actually 2 causes of the non-positive definiteness of the covariance function when the symmetrized GBO distance metric is used in Eq. 17: (i) the intrinsic non-Euclidean nature of the GBO distance; (ii) crystallographic symmetry. If crystallographic symmetry is not considered, the GBO distance is, in fact, very nearly Euclidean, at least for the sub-manifold of [100] tilt GBs. For this sub-manifold we find that the approximation

$$\Omega \approx \sqrt{\left(\frac{\omega_{AB} - \omega_{CD}}{2}\right)^2 + (\beta_A - \beta_C)^2} \tag{18}$$

produces the same result as Eq. 16 to within 0.3° for $\Omega \in [0°, 45°]$. However, even if a truly Euclidean distance function is used, the introduction of crystallographic symmetries makes the distance become non-Euclidean. The result is that even after finding the nearest symmetric positive-definite approximating covariance function, undesirable topological artifacts are produced.

To illustrate this effect, Fig. 5 shows the result of a validation test in which GB diffusivity was

[5] The GB plane reference frame is illustrated explicitly in Fig. 3(a) as the coordinate system defined by y_R and z_R. As shown there, the parameters ω_A and ω_B are the angle of rotation for the quaternion orientations q_A and q_B respectively.

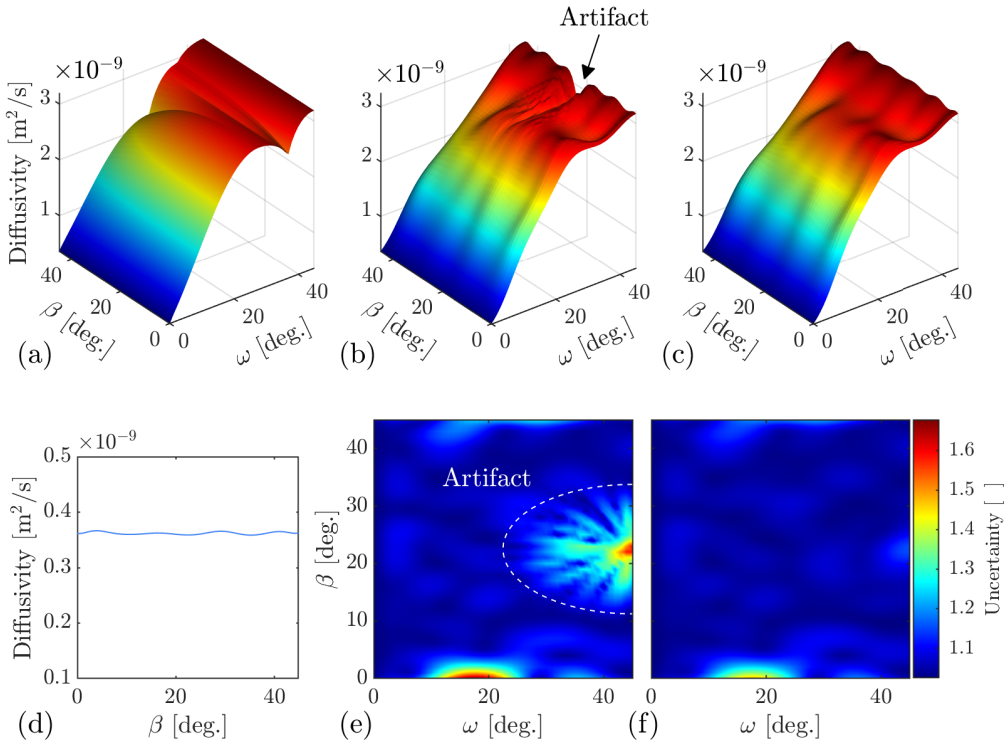

Figure 5: Comparison of inference results using different symmetrization strategies. In (a) the true (test) model is shown. This model uses the BRK model [8] to assign $\gamma(\omega, \beta)$ and then converts γ to diffusivity via the Borisov relation [4, 21]. Inference results from 100 random observation points using (b) the symmetrized distance approach, and (c) the sum-over-group-orbits approach are shown, together with their respective posterior uncertainties in (e) and (f). The uncertainty is multiplicative as described in Section 3.1. In (d) the inference result from (c) along the line $\omega = 0°$ is presented.

inferred from 100 random samples taken from a test model. Notice the spurious wrinkling artifact present in both the MAP estimate, (b), and the corresponding posterior uncertainty, (e), in the vicinity of the two-fold rotation axis.

We attribute this topological artifact to the fact that when the distance function is symmetrized it can distort the shape of the kernel in the vicinity of the symmetry axis, creating a cusp that varies with position, as shown in Fig. 6(a).

To circumvent such topological artifacts, and still enforce that $C(\mathbf{x}, \mathbf{x}')$ is argument-wise invariant under H, we take a different approach. Rather than enforcing symmetry via $d(\mathbf{x}, \mathbf{x}')$, we instead use the *unsymmetrized* GBO distance (i.e. we use Eq. 16 without minimizing over symmetrically equivalent points) and impose argument-wise invariance on $C(\mathbf{x}, \mathbf{x}')$ under H via a sum over the orbit of H as follows:

$$C_{\text{symm}}(\mathbf{x}, \mathbf{x}') = \frac{1}{|H|^2} \sum_{h \in H} \sum_{h' \in H} C(h.\mathbf{x}, h'.\mathbf{x}') \quad (19)$$

where $h, h' \in H$ are crystallographic symmetry operations, $h.\mathbf{x}$ indicates their application to the point $\mathbf{x}$, and $|H|$ is the cardinality of H. As shown in Fig. 6(b), the resulting covariance function satis-

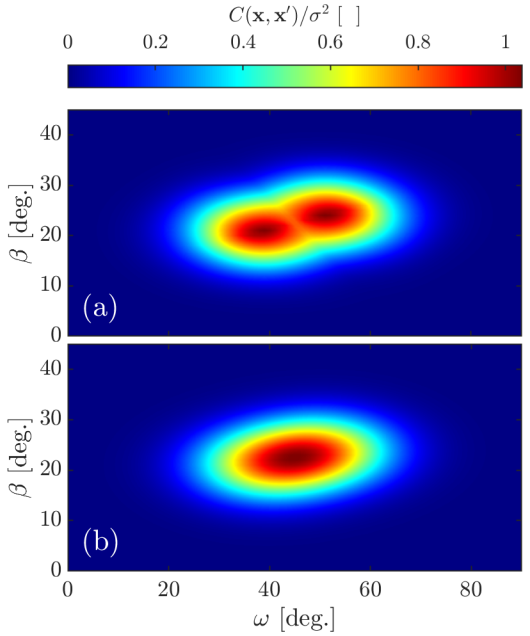

Figure 6: Symmetrized covariance function evaluated at $\mathbf{x} \equiv (\omega, \beta) = (39°, 21°)$, using (a) the symmetrized distance approach, and (b) the sum-over-group-orbits approach.

fies the 2-fold symmetry without introducing the artifact causing cusp.

Figure 5(c) and (f) show the result of the same validation test as before, but using $C_{\text{symm}}(\mathbf{x}, \mathbf{x}')$ with the unsymmetrized GBO distance metric. Notice the dramatic reduction of the spurious topological wrinkling artifact in both the MAP estimate (c) and the posterior uncertainty (f). For our calculations we still apply `nearestSPD()` to the result of Eq. 19 to make sure that neither the slight non-Euclidean nature of the unsymmetrized GBO metric, nor any numerical floating-point errors lead to violation of the constraint of symmetric positive-definiteness.

2.6.2. No-boundary Singularity

In addition to the crystallographic symmetry already discussed, there is a fundamentally different type of invariance that must also be considered: the no-boundary singularity. In the GB octonion formalism, points for which the misorientation angle, ω, is equal to zero, but having different GB normals, β, are treated as distinct GB configurations. However, all such points are physically indistinguishable, since when $\omega = 0°$ there is no GB, rather we have a single crystal. This issue is referred to as the "no-boundary singularity," and different authors have taken different approaches to handle it in the context of distance calculations [43, 45–47, 53, 54] and the definition of GB FZs [37].

In the present context, this issue manifests itself because we expect that any valid structure-property model will return the same property value for physically indistinguishable GBs even if those GBs have distinct crystallographic parameters. This means that we have the physical constraint $\mathscr{D}(0, \beta) = \mathscr{D}(0, 0)$. Unfortunately, imposing such a constraint via the covariance function is more difficult than was enforcing crystallographic symmetry, and doing so appears to be in direct competition with the constraint of positive-definiteness. Hopefully future work will find an elegant solution to this issue. In the meantime we propose, as a workaround, an approximate solution. Rather than imposing this invariance on the *covariance function*, we impose it on the *data* by augmenting our dataset with copies of no-boundary points along the line $\omega = 0°$. The results shown in Fig. 5(b) and (c) both implement this approach with the single-crystal data point, $\mathscr{D}(0, 0)$, replicated along the line $\omega = 0°$ at a resolution of 1° in β. As is apparent, the resulting inferred function along the line $\omega = 0°$ is almost constant. Figure 5(d) shows a close up view of the trace of the inferred model from Fig. 5(c) along the line $\omega = 0°$ (which corresponds to the no-boundary condition), demonstrating that while this invariance is not exactly satisfied, it is reproduced to a very good approximation.

3. Results & Discussion

3.1. Structure-Property Model Inference

We have applied the inference methods just presented in Section 2 to the data summarized in Table 1 to obtain a structure-property model for H diffusivity in Ni [100] tilt GBs at 700 K.

For the prior model, we chose a constant value of $\mathbf{m}_0 = \frac{1}{2}\left(\ln(D_{\text{max}}) + \ln(D_{\text{min}})\right) = -19.8471$, where $D_{\text{min}} = 1.3653 \times 10^{-9}$ m^2/s and $D_{\text{max}} = 4.2249 \times 10^{-9}$ m^2/s are the minimum and maximum diffusivities that were observed in the data. For the prior model covariance matrix, $\mathbf{C}_{\mathbf{m}_0}$, we used Eq. 17 with a uniform marginal uncertainty of $\sigma = \frac{1}{2}\left(\ln(D_{\text{max}}) - \ln(D_{\text{min}})\right) = 0.5648$, and we used Eq. 19 to perform symmetrization via the sum-over-group-orbits approach.

For the priors on the data, we chose $\mathbf{d}_0 = \{d_{0,1}, d_{0,2}, \ldots, d_{0,N}\}$, where $d_{0,n} = \ln\left(D_n^{\text{obs}}\right)$ represents the logarithm of the diffusivity observations. The no-boundary singularity invariance was enforced by the procedure described in Section 2.6.2. For the prior data covariance matrix, we chose $\mathbf{C}_{\mathbf{d}_0} = \text{diag}\left(\sigma^2_{d_{0,1}}, \sigma^2_{d_{0,2}}, \ldots, \sigma^2_{d_{0,N}}\right)$, where $\sigma_{d_{0,n}}$ represents the uncertainty of the n-th diffusivity observation.

For the MD diffusivity data, DS01, from Table 1, D_n^{obs} is defined as

$$D_n^{\text{obs}} = \exp\left(\frac{1}{K}\sum_{k=1}^{K}\ln(D_{n,k})\right) \tag{20}$$

where $D_{n,k}$ is the estimate of the diffusivity of the n-th GB from the k-th sub-interval in the bootstrapping procedure (see Fig. 1 and Section 2.2.1).

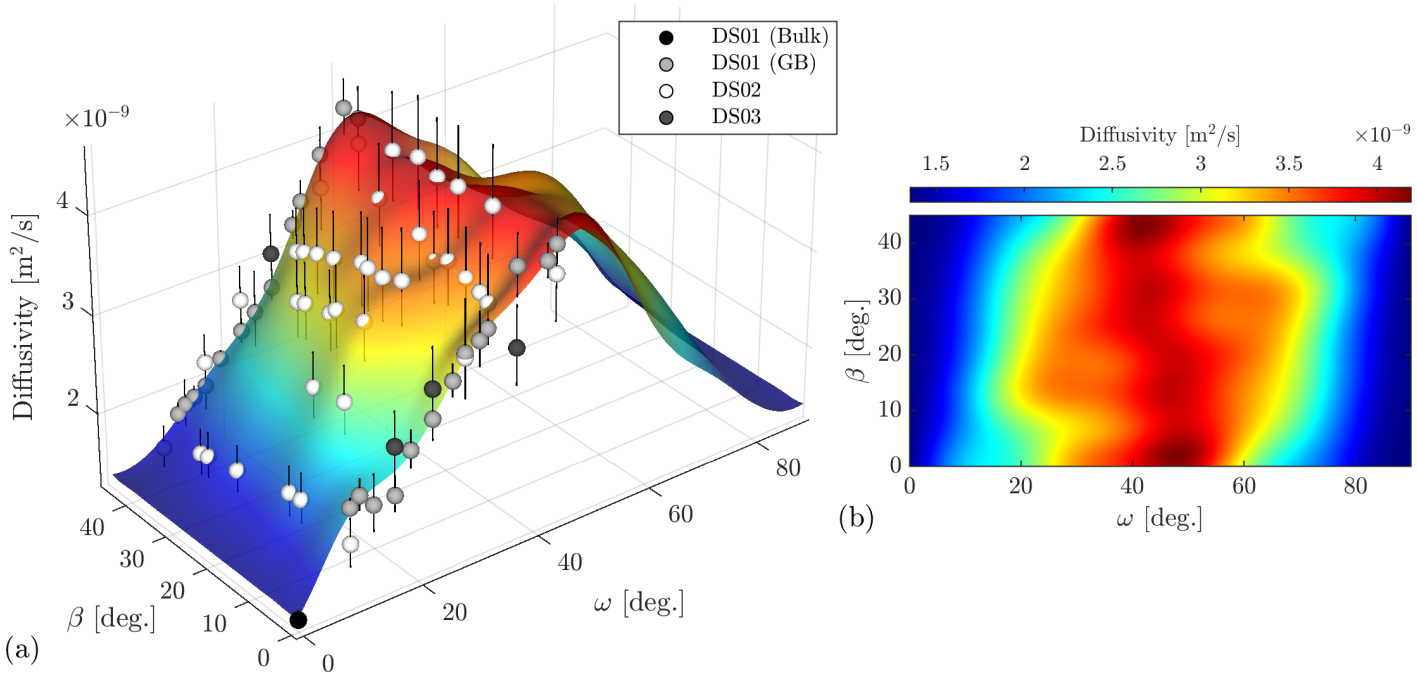

Figure 7: Colored surface in (a) and (b) shows the inferred function, $\widetilde{\mathscr{D}}(\omega, \beta)$, for H diffusivity in [100] tilt GBs in Ni at 700 K using datasets DS01, DS02, DS03. In (a) data points are included with black lines indicating the observation uncertainty for each point. In (b) a top-down view of (a) is provided showing the 2-fold symmetry of the inferred function.

The definition of D_n^{obs} provided in Eq. 20 simply makes $d_{0,n}$ equal to the mean of the logarithmic bootstrap estimates for a given GB; however, we note that this definition of D_n^{obs} also coincides with the median of the diffusivities themselves to the degree that the distribution is lognormal. The uncertainty corresponding to D_n^{obs}, is also calculated from the bootstrap estimates according to

$$\sigma_{d_{0,n}}^2 = \frac{1}{K-1}\sum_{k=1}^{K}\left[\ln(D_{n,k}) - \ln\left(D_n^{\text{obs}}\right)\right]^2 \qquad (21)$$

i.e. $\sigma_{d_{0,n}}$ is defined as the standard deviation of the logarithm of the diffusivities obtained by the bootstrapping procedure.

For the MD energy data, DS02 and DS03, from Table 1, D_n^{obs} was obtained by converting the energy data to diffusivity via the procedure described in Section 2.1 and Appendix A. The corresponding uncertainty values $\sigma_{d_{0,n}}$ were obtained via the same procedure (see Appendix A) based on the estimated uncertainty in the energy data (i.e. the uncertainty in the energy observations is propagated to the corresponding estimates of diffusivity).

Figure 7 presents our results: an inferred structure-property model, $\mathscr{D}(\omega, \beta)$, for H diffusivity in Ni [100] tilt GBs at 700 K. The numerical values of this structure-property model evaluated over a grid of $1^\circ \times 1^\circ$ resolution are provided as a `.mat` file in the supplementary material.

As is apparent from Fig. 7(b), the inferred structure-property model exhibits the two-fold symmetry that is required by crystallographic constraints, is approximately constant along the line $\omega = 0^\circ$ (i.e. it respects the no-boundary singularity invariance, at least approximately), and is consistent with the data (see Fig. 7(a)).

The marginal posterior uncertainty is calculated from the posterior covariance function (see Eq. 10) according to

$$\sigma_{\widetilde{m}}(\mathbf{x}) = \sqrt{C_{\widetilde{m}}(\mathbf{x}, \mathbf{x})}. \qquad (22)$$

One may then write, as a shorthand notation, the posterior model with its uncertainty as $\widetilde{m}(\mathbf{x}) \pm \sigma_{\widetilde{m}}(\mathbf{x})$. To write this in the desired units of diffusivity, we must invert the log-transformation to obtain $\widetilde{\mathscr{D}}(\mathbf{x}) = \exp(\widetilde{m}(\mathbf{x}))$ and $\sigma_{\widetilde{\mathscr{D}}}(\mathbf{x}) = \exp(\sigma_{\widetilde{m}}(\mathbf{x}))$. In these original units, the shorthand expression becomes $\widetilde{\mathscr{D}}(\mathbf{x}) \circledast \sigma_{\widetilde{\mathscr{D}}}(\mathbf{x})$,

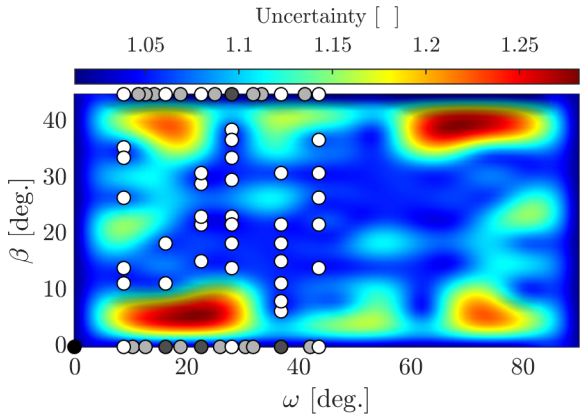

Figure 8: The multiplicative uncertainty ($\sigma_{\widetilde{\mathscr{D}}}$) of the inference is shown together with the locations of the data.

where $\divideontimes$ indicates "multiply-or-divide" instead of "plus-or-minus". This means that when the uncertainty in the inferred model is expressed in terms of $\sigma_{\widetilde{\mathscr{D}}}(\mathbf{x})$ one should remember that it is multiplicative, *not* additive. The uncertainty $\sigma_{\widetilde{\mathscr{D}}}(\mathbf{x})$ is shown in Fig. 8.

To give an intuitive understanding of the magnitude of the multiplicative uncertainty and its meaning, we note that $D \divideontimes (1+f)$ is in the interval $\left[\frac{D}{1+f},\, D(1+f)\right] \approx \left[D(1-f),\, D(1+f)\right]$, which holds to a good approximation up to about $f = 0.3$. As an example, if $\sigma_{\widetilde{\mathscr{D}}}(\mathbf{x}) = 1.1$ then the uncertainty is approximately $\pm 10\%$.

As expected, Fig. 8 shows that the uncertainty is lowest in regions where data is abundant (recall that the observation located at $(0°, 0°)$ was replicated along the line $\omega = 0°$), and highest in regions where data is absent. The median uncertainty is 1.0767 or approximately $\pm 8\%$. We also note that the uncertainty also correctly reflects the expected two-fold crystallographic symmetry and the no-boundary singularity invariance. The reduction in the uncertainty from the prior ($\exp(\sigma) = \exp(0.5648) \approx 1.76$) to the posterior ($\sigma_{\widetilde{\mathscr{D}}} \in [1.01, 1.28]$) is also notable.

As $\widetilde{m}(\omega, \beta)$ is the mean of the posterior $\sigma_m(m)$, the corresponding structure-property model $\widetilde{\mathscr{D}}(\omega, \beta)$ is analogous to a type of median of the posterior distribution $\sigma_{\mathscr{D}}(\mathscr{D})$, which provides the probability density of an arbitrary model $\mathscr{D}$ in light of our observations and priors. $\widetilde{\mathscr{D}}(\omega, \beta)$ is therefore useful because it is a representative model with high probability density. However, the true solution to the inference problem is embodied in $\sigma_{\mathscr{D}}(\mathscr{D})$ itself, rather than any particular model realization sampled from it. Because the inference method provides $\sigma_{\mathscr{D}}(\mathscr{D})$, it facilitates not only uncertainty quantification, but also uncertainty propagation. In particular, it enables us to propagate the uncertainty in the inferred structure-property model through mesoscale simulations to study the interaction between GB structure-property model uncertainty and the effective properties of GB networks in polycrystals.

3.2. Uncertainty Propagation in Mesoscale Simulations

Traditionally, mesoscale simulations rely on constitutive structure-property models that are assumed to be exactly known. In reality, no structure-property model is exact and the uncertainty in the structure-property model will produce uncertainty in the results of the mesoscale simulations. The current inference framework enables this to be considered in a rigorous way. In this section, we describe our approach and results for uncertainty propagation in mesoscale GB network diffusion simulations to study the interaction between GB structure-property model uncertainty and GB network structure.

Figure 9 provides a schematic illustration of the process. The fundamental idea is that rather than considering the structure-property model to be deterministic and exact, we consider it to be a random variable characterized by the distribution $\sigma_{\mathscr{D}}(\mathscr{D})$. We then generate samples from this distribution and perform repeated mesoscale simulations using each model realization. The result will be a distribution of simulation results (rather than a single deterministic result), whose uncertainty may be quantified.

For our implementation of this approach, we sampled $S = 100$ different models[6], $\{\mathscr{D}_s(\omega, \beta) \mid s \in [1, S]\}$, from the posterior distribution, $\sigma_{\mathscr{D}}(\mathscr{D})$ using the procedure described in Ap-

[6]To be precise, we sample models, $m_s(\omega, \beta)$, from the posterior, $\sigma_m(m)$, and then convert each $m_s(\omega, \beta)$ to $\mathscr{D}_s(\omega, \beta)$.

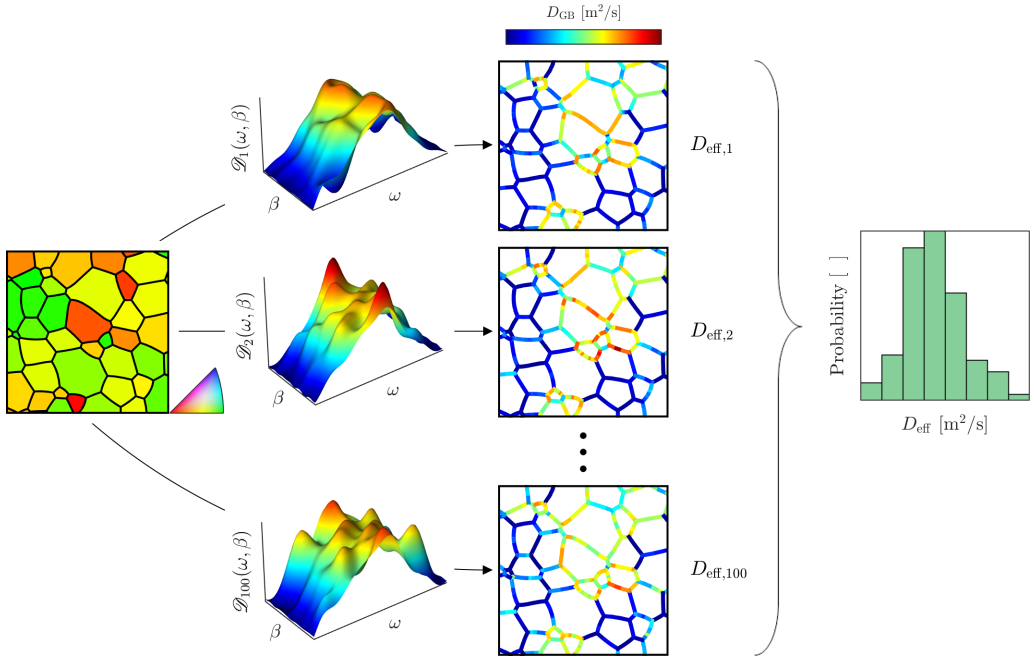

Figure 9: Schematic illustration of the uncertainty propagation process employed. For a *single* microstructure, 100 different simulations of the effective GB network diffusivity, D_{eff}, are performed. In each simulation the GB properties are assigned using a different model $\mathscr{D}_s(\omega,\beta)$. The uncertainty in the simulation results, $\sigma_{D_{\text{eff}}}$ is calculated from the resulting distribution of D_{eff}. The process shown in this figure is repeated for *each* of the 1771 microstructures considered.

pendix B. For each microstructure we assign GB properties using each of the 100 models, and calculate the effective diffusivity of the GB network, D_{eff}. The resulting distribution of D_{eff} quantifies the uncertainty in D_{eff} resulting from the uncertainty in the structure-property model $\widetilde{\mathscr{D}}(\omega,\beta)$.

To investigate the effect of the interaction between structure-property model uncertainty and GB network structure on the uncertainty of effective properties, we employed a database of 1771 two-dimensional microstructures from [35, 36], each having 100 grains, and whose triple-junction (TJ) fractions span the J-space (the space of all possible TJ fractions). TJ fractions [55–60], $\{J_i \mid i \in [0,3]\}$, denote the fraction of TJs coordinated by i low-angle GBs (LAGBs). Consequently, they quantify local spatial correlations in the GB network. We calculated the effective diffusivity of the GB network using the finite volume approach of [35, 36].

Figure 10 shows the multiplicative uncertainty of the effective diffusivity,

$$\sigma_{D_{\text{eff}}} = \exp\left(\frac{1}{S}\sum_{s=1}^{S}\left[\ln(D_{\text{eff},s}) - \langle\ln(D_{\text{eff},s})\rangle\right]^2\right) \quad (23)$$

for each of the 1771 microstructures, where $\langle\cdot\rangle$ implies the mean. The range of observed values of $\sigma_{D_{\text{eff}}}$ corresponds to approximately between $\pm 2\%$ and $\pm 5\%$, which indicates that the microstructures with the largest effective property uncertainty exhibit more than twice the uncertainty of those with the least uncertainty. There is a notable trend of increasing uncertainty from the J_1 corner towards the J_2 corner. This is visible in Fig. 10(a) as the color transitions from dark blue at the J_1 corner to orange/red near the the J_2 corner.

To view the trends more clearly we plot $\sigma_{D_{\text{eff}}}$ vs. each of the J_i coordinates in Fig. 10(b). We also quantify the correlation between $\sigma_{D_{\text{eff}}}$ and each of the J_i using Spearman's correlation coefficient, ρ_S. Like the more familiar Pearson correlation coefficient, ρ_P, Spearman's correlation coefficient ranges between $\rho_S \in [-1,1]$ and is invariant to monotone transformations of either of the variables [61]. However, whereas ρ_P measures only linear correlations, ρ_S measures any monotonic correlation—whether linear or non-linear—and is therefore more general. Because we have no *a priori* reason to assume linear correlations, we employ the more general ρ_S.

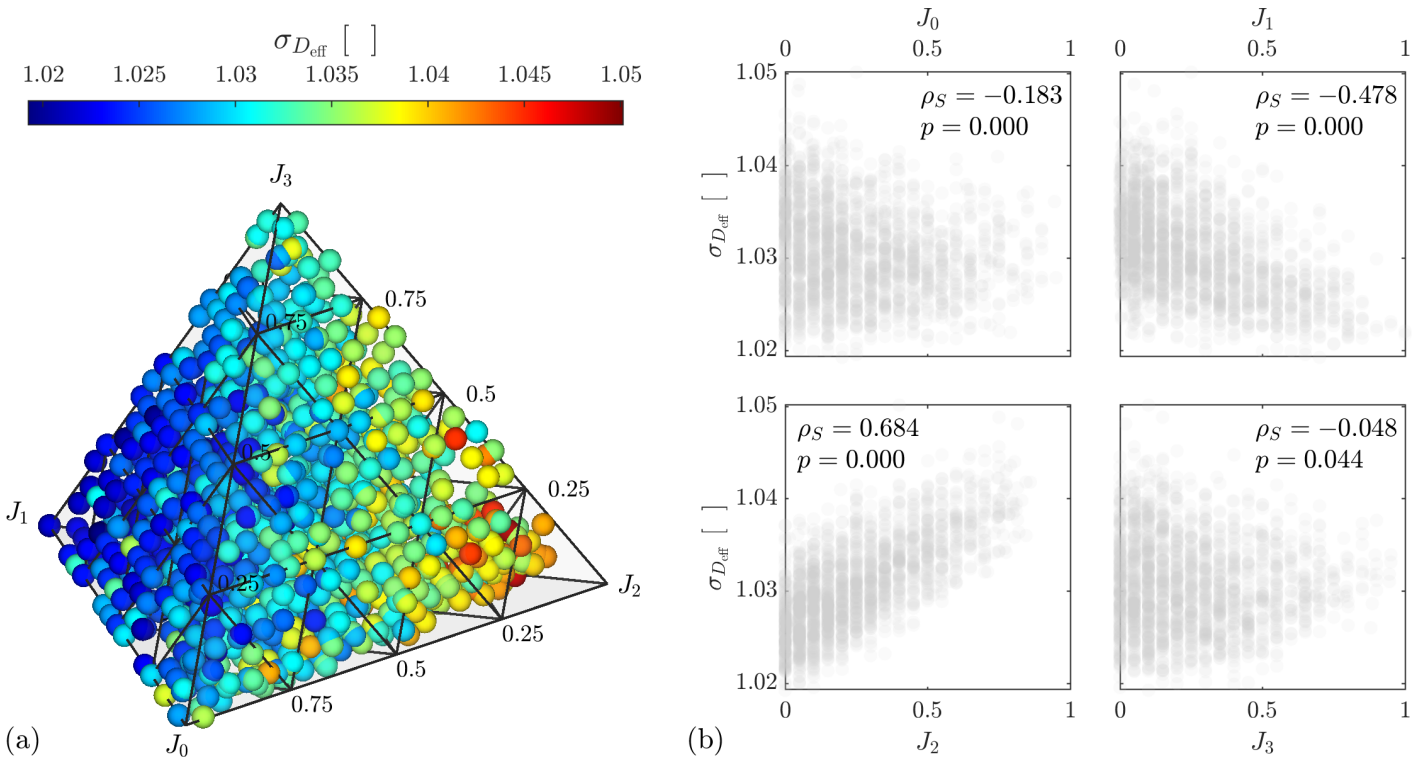

Figure 10: (a) Multiplicative uncertainty, $\sigma_{D_{\text{eff}}}$, of the effective GB network diffusivity, D_{eff}, for each of the 1771 microstructures. The range of observed uncertainty is equivalent to between roughly ±2% and ±5%, indicating that the microstructures with the largest effective property uncertainty have more than twice the uncertainty of those with the least uncertainty. Each data point corresponds to a single microstructure, and its coordinates indicate that microstructure's TJ fractions. For example, the microstructure with coordinates $[J_0, J_1, J_2, J_3] = [0.5, 0, 0.25, 0.25]$ (for which 50 % of its TJs are coordinated by 0 LAGBs, 25 % by 2 LAGBs, and 25 % by 3 LAGBs) was observed to exhibit an uncertainty of $\sigma_{D_{\text{eff}}} = 1.0364$ (i.e. approximately ±4%). In (b) we plot $\sigma_{D_{\text{eff}}}$ vs. each of the respective J_i coordinates so that the trends are easier to see. We also provide the Spearman correlation coefficient, ρ_S, and the corresponding p-value to quantify the correlation between $\sigma_{D_{\text{eff}}}$ and each of the respective J_i.

The value of ρ_S is provided together with the corresponding p-value for each plot in Fig. 10(b). We note that the p-value indicates that a statistically significant correlation between $\sigma_{D_{\text{eff}}}$ and each of the respective J_i exists at the $\alpha = 0.05$ level; however, the magnitude of that correlation (the value of ρ_S) varies widely, with $\sigma_{D_{\text{eff}}}$ having a strong positive correlation with J_2, a strong negative correlation with J_1, and very small negative correlations with J_0 and J_3.

3.3. Origin of Correlations Between $\sigma_{D_{\text{eff}}}$ and J_i

The fact that such correlations exist is significant. The TJ fractions of a microstructure directly quantify the type and magnitude of local spatial correlations in the GB network. They can be used to define a set of order parameters that characterize (i) the degree of mixing vs. segregation between LAGBs and high-angle GBs (HAGBs) and (ii) whether clusters of LAGBs (or HAGBs) form compact or extended topologies [62]. Because $\sigma_{D_{\text{eff}}}$ is a direct result of the posterior distribution of structure-property models, $\sigma_{\mathscr{D}}(\mathscr{D})$, the existence of a trend in $\sigma_{D_{\text{eff}}}$ across the J-space implies a relationship or interaction between $\sigma_{\mathscr{D}}(\mathscr{D})$ and the spatial correlations embodied in the TJ fractions.

To analyze this effect we first recall the fact, noted earlier, that because it is Gaussian, the posterior distribution over the model space, $\sigma_m(m)$, can be compactly described in terms of its mean $\widetilde{m}(\mathbf{x})$ and covariance operator $C_{\widetilde{m}}(\mathbf{x}, \mathbf{x}')$. Moreover, $C_{\widetilde{m}}(\mathbf{x}, \mathbf{x}')$ can be decomposed according to

$$C_{\widetilde{m}}(\mathbf{x}, \mathbf{x}') = \sigma_{\widetilde{m}}(\mathbf{x})\, \sigma_{\widetilde{m}}(\mathbf{x}')\, \rho_{\widetilde{m}}(\mathbf{x}, \mathbf{x}') \tag{24}$$

where $\sigma_{\widetilde{m}}(\mathbf{x})$ is the posterior marginal uncertainty defined in Eq. 22, and $\rho_{\widetilde{m}}(\mathbf{x}, \mathbf{x}') \in [-1, 1]$ is a corre-

Table 2: Definitions and levels of the factors (independent variables) used in the full factorial experimental design and analysis of variance (ANOVA). The values for the constants defining the "off" states of $\widetilde{m}(\mathbf{x})$ and $\sigma_{\widetilde{m}}(\mathbf{x})$ were chosen to be near the mean of the actual values of the respective functions.

Factor	Levels	
	"on"	"off"
$\widetilde{m}(\mathbf{x})$	actual	$\ln(2.5 \times 10^{-9}\ \mathrm{m}^2/\mathrm{s})$
$\sigma_{\widetilde{m}}(\mathbf{x})$	actual	0.0876
$\rho_{\widetilde{m}}(\mathbf{x}, \mathbf{x}')$	actual	$\delta(\mathbf{x}, \mathbf{x}')$

lation function that defines how well-correlated the values of structure-property models generated by the posterior distribution, $\sigma_m(m)$, and evaluated at points $\mathbf{x}$ and $\mathbf{x}'$ will be. In other words, if we were to generate many samples, $\{m_s(\mathbf{x}) \mid s \in [1, S]\}$, from the posterior, $\sigma_m(m)$, and then plot $m_s(\mathbf{x})$ vs. $m_s(\mathbf{x}')$ for each sampled model, $\rho_{\widetilde{m}}(\mathbf{x}, \mathbf{x}')$ would describe the observed correlation between $m_s(\mathbf{x})$ and $m_s(\mathbf{x}')$ across all s.

This decomposition of the posterior allows us to investigate the origin of the correlation between $\sigma_{D_{\text{eff}}}$ and the J_i. The J_i quantify the short-range spatial correlations between LAGBs and HAGBs in the microstructure, and are therefore independent of any structure-property model applied to the microstructure. Consequently, any trend in $\sigma_{D_{\text{eff}}}$ over the J-space must be a result of the posterior, $\sigma_m(m)$, and therefore must originate with the 3 independent functions, $\widetilde{m}(\mathbf{x})$, $\sigma_{\widetilde{m}}(\mathbf{x})$, or $\rho_{\widetilde{m}}(\mathbf{x}, \mathbf{x}')$, that uniquely define the posterior.

It may be that the observed trend over the J-space is an artifact of the form of the structure-property models $\mathscr{D}_s(\omega, \beta)$ sampled from the posterior during the uncertainty propagation process (i.e. how the diffusivity varies over (ω, β), which, though stochastic, shares characteristic features with the model $\widetilde{\mathscr{D}}(\omega, \beta)$ shown in Fig. 7). Alternatively, the trend could be the result of the marginal uncertainty $\sigma_{\widetilde{\mathscr{D}}}(\mathbf{x})$ (shown in Fig. 8), which merely indicates the sparsity of data over (ω, β) for the particular dataset employed. Or, it is possible that the similarity of properties between different GBs over (ω, β), as quantified in $\rho_{\widetilde{m}}(\mathbf{x}, \mathbf{x}')$, may be responsible for the trend.

To test these hypotheses and determine which of these 3 factors (or any interactions between them) are primarily responsible for the observed trend in effective property uncertainty over the J-space, we performed a 3-way analysis of variance (ANOVA) using a full factorial experimental design (see Fig. 11). We considered 2 levels for each of the 3 factors $\widetilde{m}(\mathbf{x})$, $\sigma_{\widetilde{m}}(\mathbf{x})$, and $\rho_{\widetilde{m}}(\mathbf{x}, \mathbf{x}')$. For each factor, the levels were essentially "on" and "off", where the respective component functions resulting from the actual posterior constituted the "on" state, and the "off" state was defined by constant functions for $\widetilde{m}(\mathbf{x})$, and $\sigma_{\widetilde{m}}(\mathbf{x})$, or an uncorrelated state for $\rho_{\widetilde{m}}(\mathbf{x}, \mathbf{x}')$—i.e. $\rho_{\widetilde{m}}(\mathbf{x}, \mathbf{x}') = \delta(\mathbf{x}, \mathbf{x}')$. Table 2 summarizes the factors (independent variables) and their levels. We considered 4 dependent (response) variables: the Spearman correlation coefficient between $\sigma_{D_{\text{eff}}}$ and each of the respective J_i, denoted by $\rho_S(J_i, \sigma_{D_{\text{eff}}})$. The 3-way ANOVA was performed for each of these 4 response variables.

For each of the 8 experimental conditions, we followed the uncertainty propagation procedure described in Section 3.2 using a modified posterior, $\hat{\sigma}_m(m)$, for which the respective levels of the 3 factors $\widetilde{m}(\mathbf{x})$, $\sigma_{\widetilde{m}}(\mathbf{x})$, and $\rho_{\widetilde{m}}(\mathbf{x}, \mathbf{x}')$ were assigned based on the experimental condition. For each experimental condition we sampled $S = 100$ models from the modified posterior, and we repeated this process 100 times (i.e. there were 100 independent replicates of every experiment).

Figure 11 shows the distribution of $\rho_S(J_i, \sigma_{D_{\text{eff}}})$ for each of the 8 experimental conditions (over all 100 replicates). Several observations are of note.

First, when all 3 factors are in the "off" state, $\rho_S(J_i, \sigma_{D_{\text{eff}}}) \approx 0\ \forall\ i$, indicating that under these conditions the trend of $\sigma_{D_{\text{eff}}}$ over the J-space is eliminated. This confirms the hypothesis that the observed trend must be in some way a result of the

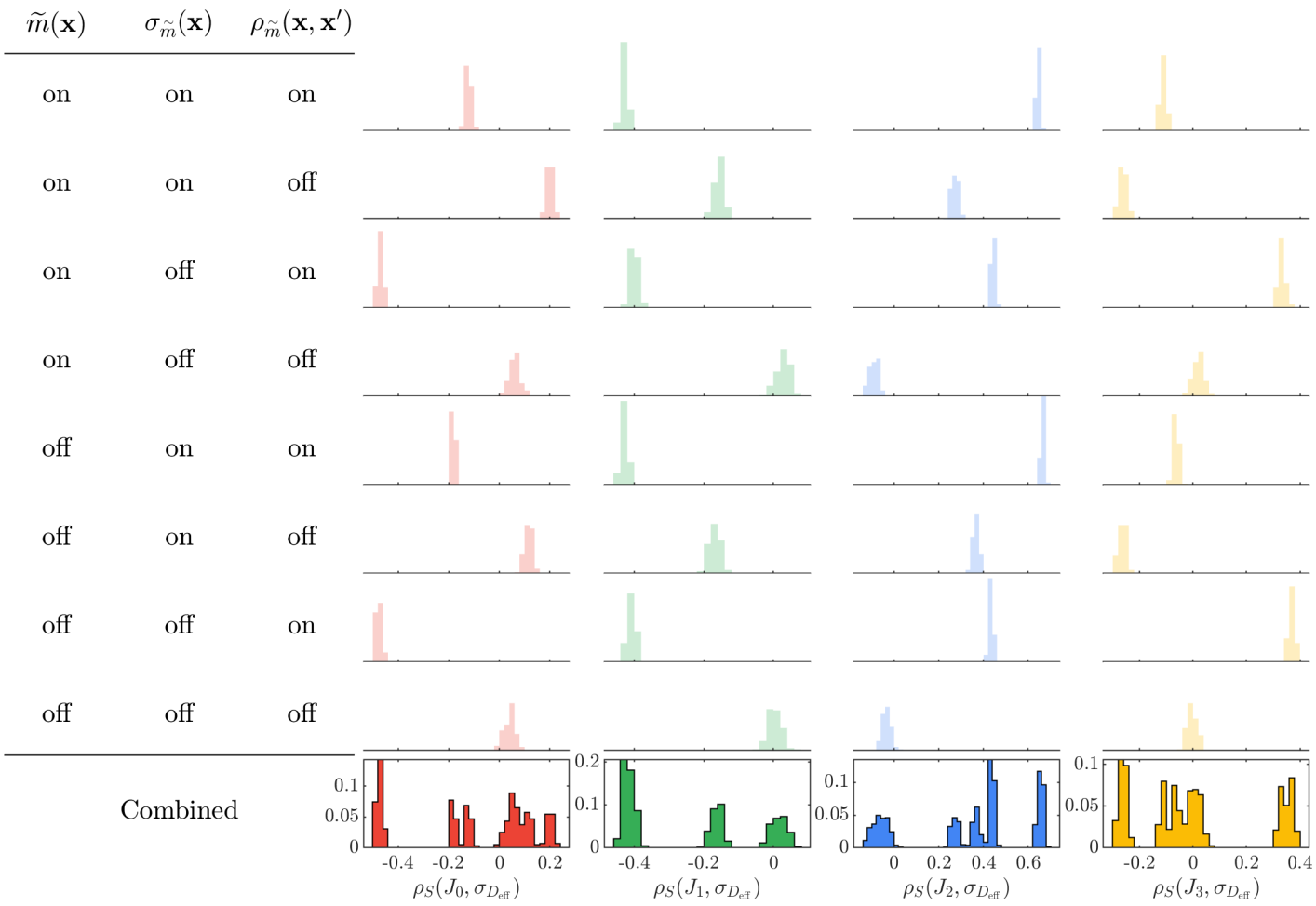

Figure 11: (left) Factor levels for each of the 8 experimental conditions. (right) Distributions of each of the respective response variables, $\rho_S(J_i, \sigma_{D_{\text{eff}}})$, over all 100 replicates for every experiment. The bottom row shows the distributions of the respective response variables when all 8 experiments are combined together. The vertical axes for all distributions correspond to probabilities and the plots without explicit vertical axes are all scaled to $[0, 0.75]$.

posterior, $\sigma_m(m)$.

Second, the distributions of the response variables fall nominally into a small number of groups that are well separated, e.g. the distributions of $\rho_S(J_1, \sigma_{D_{\text{eff}}})$ fall into 3 groups respectively centered near -0.4, -0.2, and 0. The clear separation between groups suggests that the factors that were tested explain the variation in the combined distribution (the bottom row of Fig. 11) very well.

Third, the correlations can vary widely; e.g. $\rho_S(J_2, \sigma_{D_{\text{eff}}})$ ranges from -0.14 to +0.68. This implies that changes to these factors affect not only the quantitative magnitude of the observed trend, but even its qualitative form (i.e. changes from negative correlations to positive correlations).

The ANOVA results in Table 3 provide additional insight and allow us to quantify the effect of each of the 3 factors considered, as well as other sources, such as interactions among them. The p-values indicate that all 3 factors (as well as all of their interactions) are statistically significant at the $\alpha = 0.05$ level. However, statistical significance does not tell us anything about the magnitude of the effect each factor may have—it is possible to have a very small effect that is still statistically significant.

The sum-of-squares (SS) allows us to evaluate the magnitude of the effect of each factor. The total sum-of-squares (SST) measures the total variation in each of the respective response variables, and it is simply the sum of the SS of each of the sources. The residual SS (SSR) is the amount of the SST that is

Table 3: Reduced ANOVA tables for the full factorial experimental design. The sources of variation considered, including the 3 factors $\widetilde{m}(\mathbf{x})$, $\sigma_{\widetilde{m}}(\mathbf{x})$, and $\rho_{\widetilde{m}}(\mathbf{x},\mathbf{x}')$ and their interactions, are shown in the column labeled "Source". The 4 response (dependent) variables, $\rho_S(J_i, \sigma_{D_{\text{eff}}})$ are indicated, together with the sum-of-squares (SS) and p-values for each of the sources. The total sum-of-squares (SST) is the sum of all of the SS from all sources, including the residual sum-of-squares (SSR), which is itself the amount of the SST not explained by the other sources.

	$\rho_S(J_0, \sigma_{D_{\text{eff}}})$		$\rho_S(J_1, \sigma_{D_{\text{eff}}})$		$\rho_S(J_2, \sigma_{D_{\text{eff}}})$		$\rho_S(J_3, \sigma_{D_{\text{eff}}})$	
Source	SS	p-Value	SS	p-Value	SS	p-Value	SS	p-Value
$\widetilde{m}(\mathbf{x})$	0.3754	0.0000	0.0235	0.0000	0.3272	0.0000	0.0496	0.0000
$\sigma_{\widetilde{m}}(\mathbf{x})$	9.1941	0.0000	2.0797	0.0000	17.8681	0.0000	25.1617	0.0000
$\rho_{\widetilde{m}}(\mathbf{x},\mathbf{x}')$	34.7030	0.0000	23.5479	0.0000	34.9643	0.0000	13.4270	0.0000
$\widetilde{m}(\mathbf{x}) \cdot \sigma_{\widetilde{m}}(\mathbf{x})$	0.1763	0.0000	0.0035	0.0000	0.0553	0.0000	0.0124	0.0000
$\widetilde{m}(\mathbf{x}) \cdot \rho_{\widetilde{m}}(\mathbf{x},\mathbf{x}')$	0.0203	0.0000	0.0085	0.0000	0.2377	0.0000	0.1471	0.0000
$\sigma_{\widetilde{m}}(\mathbf{x}) \cdot \rho_{\widetilde{m}}(\mathbf{x},\mathbf{x}')$	2.2796	0.0000	1.1852	0.0000	1.3356	0.0000	1.4112	0.0000
$\widetilde{m}(\mathbf{x}) \cdot \sigma_{\widetilde{m}}(\mathbf{x}) \cdot \rho_{\widetilde{m}}(\mathbf{x},\mathbf{x}')$	0.0008	0.0505	0.0001	0.4241	0.0007	0.0490	0.0016	0.0054
Residual (SSR)	0.1749		0.1593		0.1463		0.1624	
Total (SST)	46.9243		27.0079		54.9353		40.3729	

not explained by the other sources. The fact that the SSR is, in all cases, very small compared to the SST corroborates the earlier observation that the chosen factors explain the data well, and, importantly, that any variation from other sources (e.g. stochasticity or other factors that were not considered) is negligible.

The factors that have the largest effect are $\sigma_{\widetilde{m}}(\mathbf{x})$ and $\rho_{\widetilde{m}}(\mathbf{x},\mathbf{x}')$, which together account for 94 % – 96 % of the SST. Apparently, the form of the structure-property model, $\widetilde{m}(\mathbf{x})$, has a negligible effect[7], as do all of the interaction terms. Moreover, the effect of $\rho_{\widetilde{m}}(\mathbf{x},\mathbf{x}')$ was observed to be about 2 – 11 times larger than that of $\sigma_{\widetilde{m}}(\mathbf{x})$ for all of the response variables except for $\rho_S(J_3, \sigma_{D_{\text{eff}}})$, for which $\sigma_{\widetilde{m}}(\mathbf{x})$ was observed to be about 2 times larger than that of $\rho_{\widetilde{m}}(\mathbf{x},\mathbf{x}')$.

This suggests that the dominant factor that is responsible for a trend in the effective property uncertainty, $\sigma_{D_{\text{eff}}}$, over the J-space is the correlation function, $\rho_{\widetilde{m}}(\mathbf{x},\mathbf{x}')$. In other words, there is a fundamental interaction between the crystallographic correlations embodied in $\rho_{\widetilde{m}}(\mathbf{x},\mathbf{x}')$ and the spatial correlations embodied in the J_i that leads to certain types of microstructures having intrinsically larger effective property uncertainty, $\sigma_{D_{\text{eff}}}$.

It is important to note that the trend shown in Fig. 10 includes the non-negligible influence of $\sigma_{\widetilde{m}}(\mathbf{x})$. However, $\sigma_{\widetilde{m}}(\mathbf{x})$ simply measures the sparsity in our dataset, so its effect is idiosyncratic to the particular dataset. In contrast, $\rho_{\widetilde{m}}(\mathbf{x},\mathbf{x}')$ quantifies something physical about how properties of different GBs are correlated with one another and this effect is expected to be generalizable. Thus, to observe the the true underlying physical trend in $\sigma_{D_{\text{eff}}}$ over the J-space, we should consider the experimental condition when the factors $\widetilde{m}(\mathbf{x})$ and $\sigma_{\widetilde{m}}(\mathbf{x})$ are in the "off" state, and $\rho_{\widetilde{m}}(\mathbf{x},\mathbf{x}')$ is in the "on" state. This is shown in Fig. 12.

The observed trend resulting from only the impact of the natural posterior crystallographic correlations encoded in $\rho_{\widetilde{m}}(\mathbf{x},\mathbf{x}')$ has similarities with the trend previously observed in Fig. 10, which included the effects of all factors. This is to be expected since $\rho_{\widetilde{m}}(\mathbf{x},\mathbf{x}')$ was found to be the dominant factor (i.e. whether or not you include the other factors, the effect of $\rho_{\widetilde{m}}(\mathbf{x},\mathbf{x}')$ will be apparent). However, when the effect of $\rho_{\widetilde{m}}(\mathbf{x},\mathbf{x}')$ is isolated, there are also some notable differences. In particular, we find that (see, esp., Fig. 12(b))

(i) $\sigma_{D_{\text{eff}}}$ is strongly correlated with *each* of the J_i
(ii) $\sigma_{D_{\text{eff}}}$ *increases* with increasing J_2 and J_3
(iii) $\sigma_{D_{\text{eff}}}$ *decreases* with increasing J_0 and J_1.

[7] We repeated the test using a constant for the "off" state of $\widetilde{m}(\mathbf{x})$ that was many orders of magnitude different from the value given in Table 2 and the results were nearly identical.

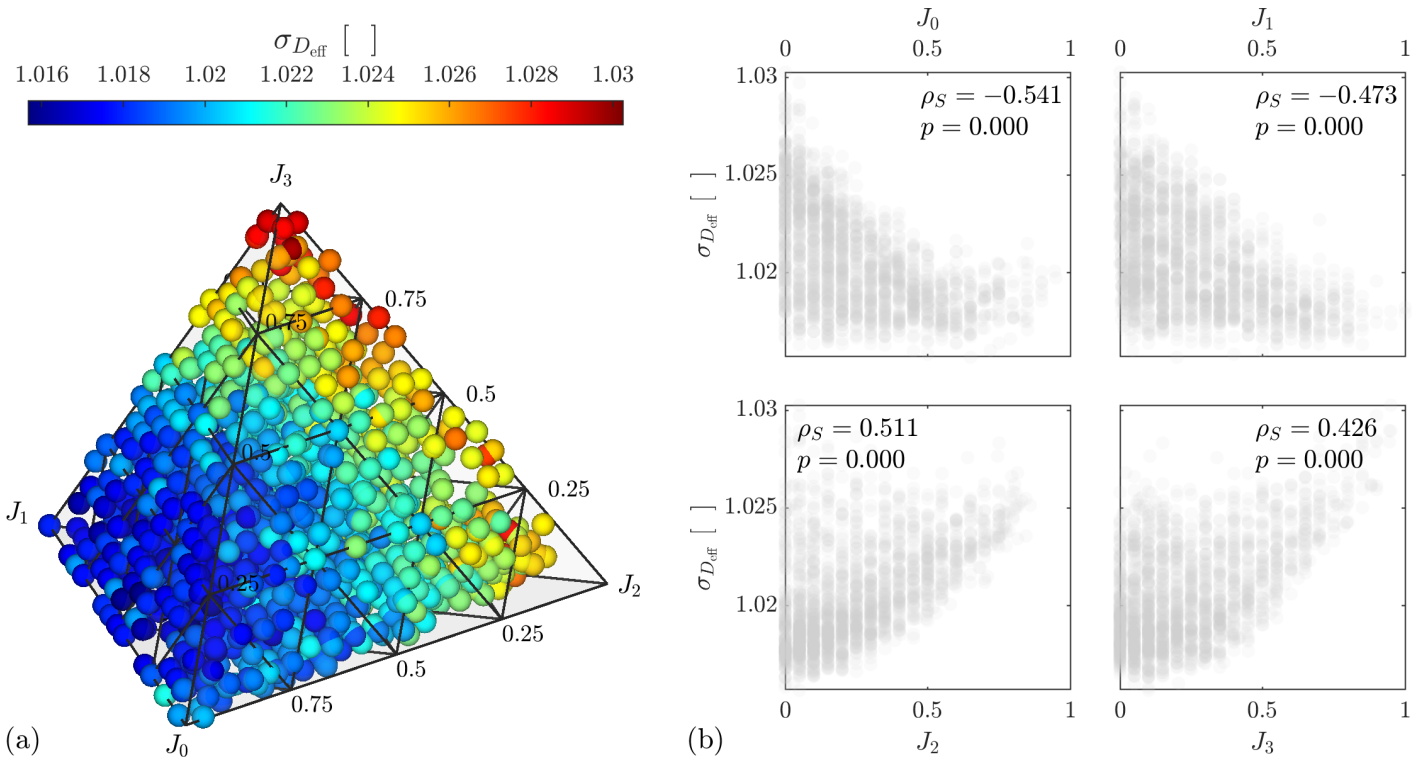

Figure 12: (a) Average multiplicative uncertainty, $\sigma_{D_{\text{eff}}}$, of the effective GB network diffusivity, D_{eff}, for each of the 1771 microstructures over all 100 replicates of the experimental condition having the factors $\widetilde{m}(\mathbf{x})$ and $\sigma_{\widetilde{m}}(\mathbf{x})$ in the "off" state, and $\rho_{\widetilde{m}}(\mathbf{x}, \mathbf{x}')$ in the "on" state. The range of observed uncertainty is equivalent to between roughly $\pm 1.6\%$ and $\pm 3\%$, indicating that the microstructures with the largest effective property uncertainty have nearly twice the uncertainty of those with the least uncertainty. In (b) we plot $\sigma_{D_{\text{eff}}}$ vs. each of the respective J_i coordinates so that the trends are easier to see. We also provide the Spearman correlation coefficient, ρ_S, and the corresponding p-value to quantify the correlation between $\sigma_{D_{\text{eff}}}$ and each of the respective J_i.

As noted earlier, one of our initial hypotheses was that the trend in $\sigma_{D_{\text{eff}}}$ might be due simply to variations in the marginal uncertainty, $\sigma_{\widetilde{m}}(\mathbf{x})$—i.e. that microstructures containing more GBs whose properties are themselves uncertain will have greater effective property uncertainty. Contrary to this expectation, these results show that the increase in effective property uncertainty with increasing J_2 and J_3 is *not* due to any particular type of GB having more intrinsic property uncertainty—since in this case all GBs have the same uncertainty because $\sigma_{\widetilde{m}}(\mathbf{x})$ was turned "off" (set to a constant)—rather it is a result of the interaction between crystallographic correlations ($\rho_{\widetilde{m}}(\mathbf{x}, \mathbf{x}')$) and spatial correlations (the J_i).

In summary, when we isolate the dominant and most physically relevant factor, $\rho_{\widetilde{m}}(\mathbf{x}, \mathbf{x}')$, we find that independent of the particular microstructures and structure-property model employed, there is a fundamental interaction between two distinct types of correlations: (i) crystallographic correlations and (ii) spatial correlations in GB networks. The result of this interaction is that certain types of microstructures (those characterized by large fractions of J_2- and/or J_3-type TJs) naturally exhibit greater uncertainty in D_{eff}.

Because the natural posterior correlations that we observe in $\rho_{\widetilde{m}}(\mathbf{x}, \mathbf{x}')$ are essentially caused by crystallographic symmetry and the assumption that GBs that are similar (with respect to the octonion distance metric) have similar properties, we anticipate that the observed trends should also be general, at least for the subset of [100] tilt GBs.

4. Conclusions and Future Outlook

In this work we presented an approach for inferring a GB structure-property model from observations of the properties and crystallography of a set

of GBs. We used this method to infer a structure-property model for H diffusivity in [100] tilt GBs in Ni at 700 K based on MD data.

We also used the inferred structure-property model to perform uncertainty propagation for mesoscale simulations of the effective diffusivity of GB networks in polycrystals. By considering a large number of diverse microstructures, we observed a fundamental interaction between crystallographic correlations and spatial correlations in GB networks that causes certain types of microstructures (those with large populations of J_2- and/or J_3-type TJs) to exhibit intrinsically larger uncertainty in their effective properties.

In addition to the major results and observations just listed, the present work contributes several important methodological developments for inference and uncertainty quantification of GB structure-property models including

- identification of topological artifacts in inference results if a covariance (kernel) function based on a symmetrized crystallographic distance is employed (Section 2.6.1)
- a method to construct a positive-definite covariance (kernel) function based on the octonion GB distance metric that avoids such topological artifacts, and which respects both crystallographic symmetries and (approximately) the no-boundary singularity (Section 2.6.1)
- a method to simultaneously perform model calibration and uncertainty propagation in predictions of GB diffusivity from observations of GB energy (Appendix A)

We also offer here some recommendations for future work, including best practices for reporting GB property data to facilitate uncertainty quantification, and suggestions for validation of the model predictions presented here.

As outlined in Section 2.2 uncertainty in GB property measurements and calculations is not universally characterized or reported; this is particularly true of epistemic uncertainties. To facilitate the development of reliable GB structure-property models, we recommend the following:

- Aleatoric uncertainties should be explicitly characterized and universally reported via a minimum of 3 – 10 replicates of every experimental measurement/computational calculation of a GB property. For computational calculations, this can often be accomplished using existing datasets with durations on the order of normal MD timescales without performing any additional simulations, using techniques like the one we demonstrate in Section 2.2.

  Admittedly, experimental replicates can be expensive and time consuming to produce. So, if replicates are not feasible, other attempts to quantify uncertainty (e.g. based on estimates of instrument resolution) may be necessary.

- There is a significant need for characterization of epistemic uncertainty in both MD simulations and experimental measurements of GB diffusivity (and other properties). For atomistic calculations, this might take several forms including (i) quantifying the distribution of errors between calculated GB properties and experimentally measured properties (and/or first-principles calculations) for a given interatomic potential; or (ii) uncertainty propagation from calibration of interatomic potential parameters to GB property predictions. See [25] for additional possible routes.

  In the case of experimental measurements, some GB properties (like GB diffusivity and GB energy) are generally obtained indirectly through fitting of an analytical theory (e.g. the Fisher model for GB diffusion profiles [63], or the Herring condition for GB energy [64]) to observed data, and therefore the theoretical (epistemic) uncertainties should be propagated to the reported measurements.

- When reporting uncertainties, it is preferable that the entire distribution of repeated measurements be published. If the distribution is instead summarized it should be done in a way that is consistent with the nature of the properties being measured. At a minimum, uncertainties in GB properties that are strictly non-negative should be reported using asymmetric

parameters such as quartiles, instead of assuming symmetric parameters like standard deviations. Moreover, when reporting summarized uncertainties in the form of such parameters, any assumptions about the form of the distribution should be validated through, e.g., tests of normality or log-normality. A convenient qualitative assessment can be easily performed using probability plots, though it is preferable to supplement these with statistical tests.

For validation and improvement of the structure-property model presented here, or other structure-property models developed in the future, we recommend the following:

- While the prediction of GB diffusivity from GB energy via the process presented here is convenient and explicitly includes the propagation of uncertainties (see Appendix A), the resulting uncertainties are much larger than for direct calculations/measurements of GB diffusivity (see Fig. 14) as they incorporate uncertainty from both the calculated GB energies and the conversion process itself (i.e. the theoretical uncertainty). It would therefore be beneficial to expand the existing dataset and reduce its uncertainty by directly calculating GB diffusivities for a much broader survey of GB types spanning the full 5D space of GB characters. This will be true for other GB properties as well.

- Characterization of the distribution of property values across the metastable states of GBs will enable more rigorous uncertainty quantification and more accurate mesoscale modeling. In the most general sense, a GB structure-property model should be a *distribution* of property values for every GB—i.e. a model of the form $f(P, \omega, \theta, \phi, \alpha, \beta)$ where $f(\cdot)$ is a distribution, P is the property of interest, and $\{\omega, \theta, \phi, \alpha, \beta\}$ are the crystallographic parameters of a GB. However, given the current state of the art, this is admittedly quite challenging, so in the short term, it would be helpful to at least estimate, for a small number of GBs in the system of interest, to what degree the properties of a single GB state (e.g. the 0 K or finite temperature minimum energy state) are representative (i.e. how tight the distribution of properties is across metastable states).

Acknowledgement

The material presented here is based upon work supported by the National Science Foundation under Grant No. 1610077. We gratefully acknowledge the use of high-performance (supercomputing) computational resources provided by Brigham Young University's Office of Research Computing.

A. Converting GB Energy to Diffusivity

To convert GB energy to diffusivity, we couple the Borisov relation [4, 21] with a Bayesian approach to simultaneous model calibration and prediction. The Borisov relation can be expressed as

$$\ln \mathbf{D}_{\mathrm{GB}} = \alpha\left(\frac{\boldsymbol{\gamma}}{kT}\right) + \ln D_{\mathrm{B}} + \ln \lambda \tag{25}$$

where $\mathbf{D}_{\mathrm{GB}} = [D_{\mathrm{GB},1}, D_{\mathrm{GB},2}, \ldots, D_{\mathrm{GB},N}]^{\intercal}$, and $\boldsymbol{\gamma} = [\gamma_1, \gamma_2, \ldots, \gamma_N]^{\intercal}$ are vectors of observations of GB diffusivity and energy, respectively, and all other parameters—D_B (bulk diffusivity), α, and λ—are scalars.

As discussed in Section 2.2, the fact that both diffusivity and energy are in $\mathbb{R} \geq 0$ implies that prior models for their uncertainties cannot be normally distributed; however, Gaussian priors can be used if a logarithmic change of variables is employed. The physical interpretation of α and λ suggests that these parameters are also non-negative (see [4] for an explanation of these parameters). Consequently, we recast Eq. 25 as

$$\mathbf{y} - \frac{1}{kT} e^{\mathbf{x}+u} - z - v = \mathbf{0} \tag{26}$$

where

$$\mathbf{x} = \ln \boldsymbol{\gamma} \tag{27}$$
$$\mathbf{y} = \ln \mathbf{D}_{\mathrm{GB}} \tag{28}$$
$$z = \ln D_{\mathrm{B}} \tag{29}$$
$$u = \ln \alpha \tag{30}$$
$$v = \ln \lambda \tag{31}$$

Our *a priori* assumptions on the parameters are that they are uncorrelated and Gaussian. The joint prior distribution is then a Gaussian with mean

$$\mathbf{m}_0 = [\mathbf{x}_0, \mathbf{y}_0, z_0, u_0, v_0]^\intercal \tag{32}$$

and covariance matrix

$$\mathbf{C}_{\mathbf{m}_0} = \mathrm{diag}\big(\sigma^2_{x_1}, \ldots, \sigma^2_{x_N}, \sigma^2_{y_1}, \ldots, \sigma^2_{y_N}, \sigma^2_z, \sigma^2_u, \sigma^2_v\big) \tag{33}$$

Similar to the approach of [14] (see pg. 273), let us define a vector $\mathbf{d}$ such that

$$\mathbf{d} = g\left(\mathbf{m}\right) = \mathbf{y} - \frac{1}{kT} e^{\mathbf{x}+u} - z - v \tag{34}$$

Our *a priori* assumptions on the Borisov relation itself are that it holds only approximately, and that uncertainties are again Gaussian. Consequently, the prior distribution on $\mathbf{d}$ will be a Gaussian with mean $\mathbf{d}_0 = \mathbf{0}$ and diagonal covariance matrix $\mathbf{C}_{\mathbf{d}_0}$.

With these priors, the posterior distribution will be of the form

$$\sigma_m(\mathbf{m}) \propto \exp\left(-\frac{1}{2} S(\mathbf{m})\right) \tag{35}$$

where

$$\begin{aligned} S(\mathbf{m}) = {} & (g(\mathbf{m}) - \mathbf{d}_0)^\intercal \mathbf{C}^{-1}_{\mathbf{d}_0} (g(\mathbf{m}) - \mathbf{d}_0) \\ & + (\mathbf{m} - \mathbf{m}_0)^\intercal \mathbf{C}^{-1}_{\mathbf{m}_0} (\mathbf{m} - \mathbf{m}_0) \end{aligned} \tag{36}$$

The non-linearity of Eq. 26 implies that the posterior is not strictly Gaussian, but we may expect it to be approximately Gaussian near the MAP point, $\widetilde{\mathbf{m}}$.

We wish to predict GB diffusivity from GB energy via the Borisov relation, while explicitly incorporating uncertainties in our observations of both GB diffusivity and GB energy as well as uncertainty in the Borisov relation itself. We do this by performing the inference in 2 passes.

In the first pass, we let

$$\mathbf{x}_0 = \mathbf{x}_{\mathrm{obs}} \tag{37}$$

$$\mathbf{y}_0 = \mathbf{y}_{\mathrm{obs}} \tag{38}$$

where $\mathbf{x}_{\mathrm{obs}}$ and $\mathbf{y}_{\mathrm{obs}}$ are vectors that correspond to pairs of observations of GB energy and diffusivity respectively (we use dataset DS01 from Table 1 together with the corresponding 0 K GB energies reported in our prior work [4]). We also set $\mathbf{C}_{\mathbf{d}_0} = \mathbf{0}$, which forces the posterior estimates $\widetilde{\mathbf{x}}$ and $\widetilde{\mathbf{y}}$ to fall on the line $g(\widetilde{\mathbf{m}}) = \mathbf{d}_0$. We then obtain the MAP point, $\widetilde{\mathbf{m}}$ via the following iterative steepest-descent algorithm (see Eq. 6.308 of [14])

$$\begin{aligned} \widetilde{\mathbf{m}}_{k+1} = \widetilde{\mathbf{m}}_k - \nu_k \Big(& \mathbf{C}_{\mathbf{m}_0} \mathbf{G}_k^\intercal (\mathbf{C}_{\mathbf{d}_0} + \mathbf{G}_k \mathbf{C}_{\mathbf{m}_0} \mathbf{G}_k^\intercal)^{-1} \\ & \cdot (\mathbf{d}_0 - g(\widetilde{\mathbf{m}}_k) + \mathbf{G}_k (\widetilde{\mathbf{m}}_k - \mathbf{m}_0)) \Big) \end{aligned} \tag{39}$$

with the variable step-size, ν_k, given by

$$\nu_k = \frac{\mathbf{a}_k^\intercal \mathbf{C}^{-1}_{\mathbf{m}_0} \mathbf{a}_k}{\mathbf{a}_k^\intercal \mathbf{C}^{-1}_{\mathbf{m}_0} \mathbf{a}_k + \mathbf{b}_k^\intercal \mathbf{C}^{-1}_{\mathbf{d}_0} \mathbf{b}_k} \tag{40}$$

and where

$$\mathbf{a}_k = \mathbf{C}_{\mathbf{m}_0} \mathbf{G}_k^\intercal \mathbf{C}^{-1}_{\mathbf{d}_0} \left(g(\widetilde{\mathbf{m}}_k) - \mathbf{d}_0\right) + (\widetilde{\mathbf{m}}_k - \mathbf{m}_0) \tag{41}$$

$$\mathbf{b}_k = \mathbf{G}_k \mathbf{a}_k \tag{42}$$

Using the posterior estimates of $\widetilde{u}$ and $\widetilde{v}$ from this first pass, together with the prior data, $\mathbf{x}_0$, $\mathbf{y}_0$, and z_0, we obtain an estimate of the modelization uncertainty from the distribution of the residuals, which can be expressed as

$$\boldsymbol{\delta}_0 = \mathbf{y}_0 - \frac{1}{kT} e^{\mathbf{x}_0 + \widetilde{u}} - z_0 - \widetilde{v} \tag{43}$$

as shown in Fig. 13.

Using the dispersion of the residuals, we perform the second pass with $\mathbf{C}_{\mathbf{d}_0} = \sigma^2_{\boldsymbol{\delta}_0}$ and let

$$\mathbf{x}_0 = [\mathbf{x}_{\mathrm{obs}}, \mathbf{x}_{\mathrm{query}}]^\intercal \tag{44}$$

$$\mathbf{y}_0 = [\mathbf{y}_{\mathrm{obs}}, \mathbf{y}_{\mathrm{pred}}]^\intercal \tag{45}$$

where $\mathbf{x}_{\mathrm{query}}$ represents the set of GB energy observations (DS02 and DS03) for which we wish to predict the corresponding GB diffusivity, as embodied in $\mathbf{y}_{\mathrm{pred}}$.

We choose our prior estimates of $\mathbf{y}_{\mathrm{pred}}$, u_0, and v_0 arbitrarily to be equal to zero, and we set the corresponding uncertainties to an arbitrarily large value (we use 10^6) representing our complete ignorance about the true value of these parameters. We then

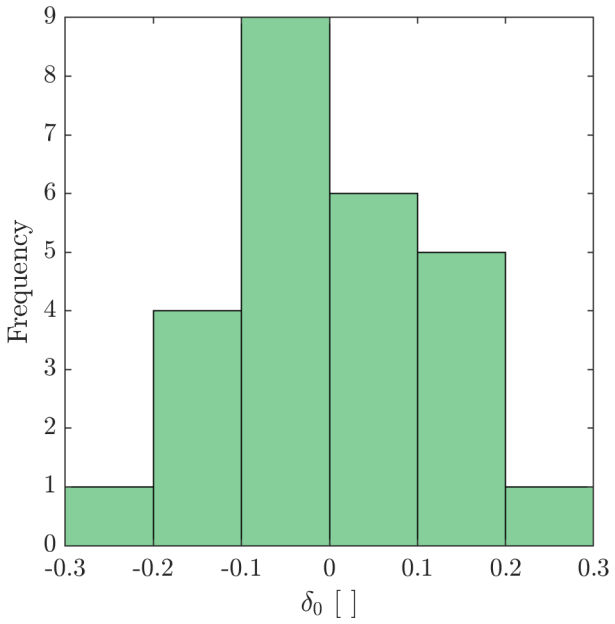

Figure 13: Distribution of the residuals (Eq. 43) of the first-pass of the model calibration procedure.

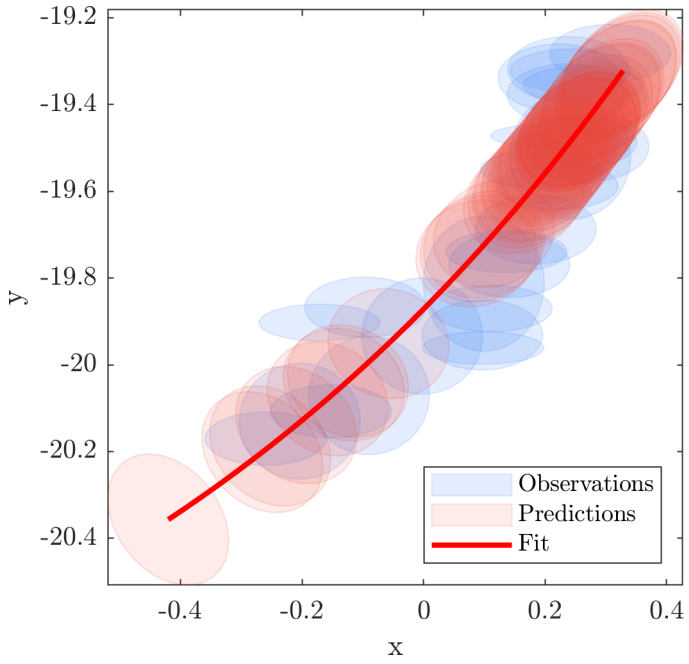

Figure 14: Bayesian predictions of GB diffusivity ($y = \ln D_{\mathrm{GB}}$) with quantified uncertainty, for specific observations of GB energy ($x = \ln \gamma$) based on calibration of the Borisov relation (Eq. 25). The one standard deviation ellipses represent the posterior uncertainty of the predictions (red ellipses) and the prior uncertainty of the calibration data (blue ellipses), respectively. The calibrated fit of the Borisov relation is also shown (red line).

perform simultaneous model calibration and prediction by again obtaining the MAP point, which we will denote $\widetilde{\widetilde{\mathbf{m}}}$, to distinguish it from the result of the first pass. We find it convenient this time to use a Newton algorithm (see Eq. 6.319 of [14]):

$$\widetilde{\widetilde{\mathbf{m}}}_{k+1} = \widetilde{\widetilde{\mathbf{m}}}_k - \left(\mathbf{C}_{\mathbf{m}_0}^{-1} + \mathbf{G}_k^{\intercal}\mathbf{C}_{\mathbf{d}_0}^{-1}\mathbf{G}_k\right)^{-1} \cdot \left(\mathbf{G}_k^{\intercal}\mathbf{C}_{\mathbf{d}_0}^{-1}\left(g(\widetilde{\widetilde{\mathbf{m}}}_k) - \mathbf{d}_0\right) + \mathbf{C}_{\mathbf{m}_0}^{-1}\left(\widetilde{\widetilde{\mathbf{m}}}_k - \mathbf{m}_0\right)\right) \tag{46}$$

The result of this two step procedure is shown in Fig. 14 where the red ellipses provide the desired predictions of GB diffusivity with quantified uncertainty, for specific observations of GB energy (which themselves have uncertainty). These predictions incorporate both the uncertainty of the observations, calibration data, and the Borisov relation itself.

B. Sampling the Posterior

As we demonstrate in Section 3.2, one major advantage of the Bayesian approach is that it facilitates uncertainty propagation for mesoscale simulations. To realize this potential it is necessary to generate samples from the posterior. In general, this is done using Monte Carlo methods [14, 34]. However, when the posterior is Gaussian, this can be accomplished in a more efficient way [65] as follows.

Consider a discretization of the model domain into N points. Generate N independent samples, z_n, from a standard normal distribution (i.e. $z_i \sim \mathcal{N}(0,1)$. Then $\mathbf{z} = [z_1, z_2, \ldots, z_N] \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ is a sample from the standard multivariate normal distribution. Finally, form $\mathbf{m} = \widetilde{\mathbf{m}} + \mathbf{z}^{\intercal}\mathbf{U}$, where $\mathbf{U}$ is the upper-triangular Cholesky factor of the target covariance matrix $\mathbf{C}_{\widetilde{\mathbf{m}}}$. The result is that $\mathbf{m} \sim \mathcal{N}(\widetilde{\mathbf{m}}, \mathbf{C}_{\widetilde{\mathbf{m}}})$ is a sample from the multivariate normal distribution with mean $\widetilde{\mathbf{m}}$ and covariance $\mathbf{C}_{\widetilde{\mathbf{m}}}$.

This can be accomplished readily in MATLAB via the built-in `mvnrnd()` function.

References

[1] 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.

[2] D. L. Olmsted, E. A. Holm, S. M. Foiles, Survey of computed grain boundary properties in face-centered cubic metals—II: Grain boundary mobility, Acta Materialia 57 (13) (2009) 3704–3713. doi:10.1016/j.actamat.2009.04.015.

[3] G. S. Rohrer, D. M. Saylor, B. E. Dasher, B. L. Adams, A. D. Rollett, P. Wynblatt, The distribution of internal interfaces in polycrystals, Zeitschrift für Metallkunde 95 (4) (2004) 197–214. doi:10.3139/146.017934.

[4] 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 (2021) 116882doi:10.1016/j.actamat.2021.116882.

[5] H. C. Erickson, E. R. Homer, Insights into grain boundary energy structure-property relationships by examining computed [1 0 0] disorientation axis grain boundaries in Nickel, Scripta Materialia 185 (2020) 165–169. doi:10.1016/j.scriptamat.2020.03.062.

[6] H. C. Erickson, J. Priedeman, E. Homer, [100] disorientation axis grain boundary energies and structures, Mendeley Data (2020). doi:10.17632/bz46sgxm8g.1.

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

[8] 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.

[9] K. Aki, W. H. K. Lee, Determination of three-dimensional velocity anomalies under a seismic array using first P arrival times from local earthquakes: 1. A homogeneous initial model, Journal of Geophysical Research 81 (23) (1976) 4381–4399. doi:10.1029/JB081i023p04381.

[10] K. Aki, A. Christoffersson, E. S. Husebye, Determination of the three-dimensional seismic structure of the lithosphere, Journal of Geophysical Research 82 (2) (1977) 277–296. doi:10.1029/JB082i002p00277.

[11] A. Nercessian, A. Hirn, A. Tarantola, Three-dimensional seismic transmission prospecting of the Mont Dore volcano, France, Geophysical Journal of the Royal Astronomical Society 76 (2) (1984) 307–315. doi:10.1111/j.1365-246X.1984.tb05048.x.

[12] K. Mosegaard, A. Tarantola, Probabilistic approach to inverse problems, in: International Geophysics, Vol. 81 of International Geophysics, Elsevier, 2002, pp. 237–265. doi:10.1016/S0074-6142(02)80219-4.

[13] A. Tarantola, B. Valette, Generalized nonlinear inverse problems solved using the least squares criterion, Reviews of Geophysics 20 (2) (1982) 219–232. doi:10.1029/RG020i002p00219.

[14] A. Tarantola, Inverse Problem Theory and Methods for Model Parameter Estimation, Society for Industrial and Applied Mathematics, 2005. doi:10.1137/1.9780898717921.

[15] A. Tarantola, A. Nercessian, Three-dimensional inversion without blocks, Geophysical Journal of the Royal Astronomical Society 76 (2) (1984) 299–306. doi:10.1111/j.1365-246X.1984.tb05047.x.

[16] A. Ray, D. Myer, Bayesian geophysical inversion with trans-dimensional Gaussian process machine learning, Geophysical Journal International 217 (3) (2019) 1706–1726. doi:10.1093/gji/ggz111.

[17] T. T. Hines, E. A. Hetland, Revealing transient strain in geodetic data with Gaussian process regression, Geophysical Journal International 212 (3) (2018) 2116–2130. doi:10.1093/gji/ggx525.

[18] E. Bélisle, Z. Huang, A. Gheribi, Scalable Gaussian Process Regression for Prediction of Material Properties, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), Vol. 8506 LNCS, Springer Verlag, 2014, pp. 38–49. doi:10.1007/978-3-319-08608-8_4.

[19] K. S. Reddy, S. Somasundharam, An Inverse Method for Simultaneous Estimation of Thermal Properties of Orthotropic Materials using Gaussian Process Regression, Journal of Physics: Conference Series 745 (3) (2016) 032090. doi:10.1088/1742-6596/745/3/032090.

[20] M. M. Noack, G. S. Doerk, R. Li, J. K. Streit, R. A. Vaia, K. G. Yager, M. Fukuto, Autonomous materials discovery driven by Gaussian process regression with inhomogeneous measurement noise and anisotropic kernels, Scientific Reports 10 (1) (2020) 17663. doi:10.1038/s41598-020-74394-1.

[21] V. T. Borisov, V. M. Golikov, G. V. Scherbedinskiy, Relationship between diffusion coefficients and grain boundary energy, Phys. Met. Metallogr. 17 (1964) 80–84.

[22] J. Pelleg, On the relation between diffusion coefficients and grain boundary energy, Philosophical Magazine 14 (129) (1966) 595–601. doi:10.1080/14786436608211954.

[23] M. M. Moghadam, J. M. Rickman, M. P. Harmer, H. M. Chan, The role of boundary variability in polycrystalline grain-boundary diffusion, Journal of Applied Physics 117 (4) (2015) 045311. doi:10.1063/1.4906778.

[24] R. W. Balluffi, S. M. Allen, W. C. Carter, Kinetics of Materials, John Wiley & Sons, Inc., Hoboken, NJ, USA, 2005. doi:10.1002/0471749311.

[25] X. Zhou, S. M. Foiles, Uncertainty Quantification and Reduction of Molecular Dynamics Models, in: Uncertainty Quantification and Model Calibration, InTech, 2017. doi:10.5772/intechopen.68507.

[26] G. M. Poletaev, I. V. Zorya, D. V. Novoselova, M. D. Starostenkov, Molecular dynamics simulation of hydrogen atom diffusion in crystal lattice of fcc metals, International Journal of Materials Research 108 (10) (2017) 785–790. doi:10.3139/146.111556.

[27] S. Ifergane, E. Sabatani, B. Carmeli, Z. Barkay, V. Ezersky, O. Beeri, N. Eliaz, Hydrogen diffusivity measurement and microstructural characterization of Custom 465 stainless steel, Electrochimica Acta 178 (2015) 494–503. doi:10.1016/j.electacta.2015.08.016.

[28] T. Tsuru, R. Latanision, Grain boundary transport of hydrogen in nickel, Scripta Metallurgica 16 (5) (1982) 575–578. doi:10.1016/0036-9748(82)90273-3.

[29] P. Kumar, R. Balasubramaniam, Determination of hydrogen diffusivity in austenitic stainless steels by subscale microhardness profiling, Journal of Alloys and Compounds 255 (1-2) (1997) 130–134. doi:10.1016/S0925-8388(96)02846-0.

[30] P. Bruzzoni, R. Carranza, J. Collet Lacoste, E. Crespo, Hydrogen diffusion in α-iron studied using an electrochemical permeation transfer function, Electrochimica Acta 44 (16) (1999) 2693–2704. doi:10.1016/S0013-4686(98)00406-X.

[31] A. Ksibi, E. Bourasseau, X. Iltis, D. Drouan, M. Gaudet, A. Germain, A. Pena, G. Lapertot, J.-P. Brison, R. Belin, Experimental and numerical assessment of grain boundary energies in polycrystalline uranium dioxide, Journal of the European Ceramic Society

40 (12) (2020) 4191–4201. doi:10.1016/j.jeurceramsoc.2020.04.041.

[32] 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.

[33] A. Tarantola, K. Mosegaard, Mathematical basis for physical inference (2000). arXiv:math-ph/0009029.

[34] K. Mosegaard, A. Tarantola, Monte Carlo sampling of solutions to inverse problems, Journal of Geophysical Research: Solid Earth 100 (B7) (1995) 12431–12447. doi:10.1029/94JB03097.

[35] B. D. Snow, S. G. Baird, C. Kurniawan, D. T. Fullwood, E. R. Homer, O. K. Johnson, Grain boundary structure-property model inference using polycrystals: The underdetermined case, Acta Materialia 209 (2021) 116769. doi:10.1016/j.actamat.2021.116769.

[36] C. Kurniawan, S. Baird, D. T. Fullwood, E. R. Homer, O. K. Johnson, Grain boundary structure–property model inference using polycrystals: the overdetermined case, Journal of Materials Science 55 (4) (2020) 1562–1576. doi:10.1007/s10853-019-04125-z.

[37] 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.

[38] 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.

[39] D. Ginsbourger, X. Bay, O. Roustant, L. Carraro, Argumentwise invariant kernels for the approximation of invariant functions, Annales de la faculté des sciences de Toulouse Mathématiques 21 (3) (2012) 501–527. doi:10.5802/afst.1343.

[40] D. Ginsbourger, O. Roustant, N. Durrande, On degeneracy and invariances of random fields paths with applications in Gaussian process modelling, Journal of Statistical Planning and Inference 170 (2016) 117–128. doi:10.1016/j.jspi.2015.10.002.

[41] D. Duvenaud, Automatic model construction with Gaussian processes, Doctoral thesis, University of Cambridge (2014). doi:10.17863/CAM.14087.

[42] 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. doi:10.1016/S1359-6454(00)00126-9.

[43] D. L. Olmsted, A new class of metrics for the macroscopic crystallographic space of grain boundaries, Acta Materialia 57 (9) (2009) 2793–2799. doi:10.1016/j.actamat.2009.02.030.

[44] A. P. Sutton, E. P. Banks, A. R. Warwick, The five-dimensional parameter space of grain boundaries, Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences 471 (2181) (2015) 20150442. doi:10.1098/rspa.2015.0442.

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

[46] A. Morawiec, A New Metric for the Space of Macroscopic Parameters of Grain Interfaces, Metallurgical and Materials Transactions A 50 (9) (2019) 4012–4015. doi:10.1007/s11661-019-05361-3.

[47] A. Morawiec, On distances between grain interfaces in macroscopic parameter space, Acta

Materialia 181 (2019) 399–407. doi:10.1016/j.actamat.2019.09.032.

[48] F. C. Curriero, On the Use of Non-Euclidean Distance Measures in Geostatistics, Mathematical Geology 38 (8) (2007) 907–926. doi:10.1007/s11004-006-9055-7.

[49] J. B. Boisvert, Geostatistics with Locally Varying Anisotropy, Doctoral, University of Alberta (2010). doi:10.7939/R31X5C.

[50] J. B. Boisvert, C. V. Deutsch, Programs for kriging and sequential Gaussian simulation with locally varying anisotropy using non-Euclidean distances, Computers & Geosciences 37 (4) (2011) 495–510. doi:10.1016/j.cageo.2010.03.021.

[51] N. J. Higham, Computing a nearest symmetric positive semidefinite matrix, Linear Algebra and its Applications 103 (C) (1988) 103–118. doi:10.1016/0024-3795(88)90223-6.

[52] J. D'Errico, nearestSPD (2020).

[53] A. Morawiec, Models of uniformity for grain boundary distributions, Journal of Applied Crystallography 42 (5) (2009) 783–792. doi:10.1107/S0021889809025461.

[54] J. W. Cahn, J. E. Taylor, Metrics, measures, and parametrizations for grain boundaries: a dialog, Journal of Materials Science 41 (23) (2006) 7669–7674. doi:10.1007/s10853-006-0592-8.

[55] M. E. Frary, C. A. Schuh, Grain boundary networks: Scaling laws, preferred cluster structure, and their implications for grain boundary engineering, Acta Materialia 53 (16) (2005) 4323–4335. doi:10.1016/j.actamat.2005.05.030.

[56] V. Y. Gertsman, K. Tangri, Computer simulation study of grain boundary and triple junction distributions in microstructures formed by multiple twinning, Acta Metallurgica et Materialia 43 (6) (1995) 2317–2324. doi:10.1016/0956-7151(94)00422-6.

[57] O. K. Johnson, C. A. Schuh, The uncorrelated triple junction distribution function: Towards grain boundary network design, Acta Materialia 61 (8) (2013) 2863–2873. doi:10.1016/j.actamat.2013.01.025.

[58] P. Fortier, W. Miller, K. Aust, Triple junction and grain boundary character distributions in metallic materials, Acta Materialia 45 (8) (1997) 3459–3467. doi:10.1016/S1359-6454(97)00004-9.

[59] Y. Yi, J. Kim, Characterization methods of grain boundary and triple junction distributions, Scripta Materialia 50 (6) (2004) 855–859. doi:10.1016/j.scriptamat.2003.12.010.

[60] P. Davies, V. Randle, G. Watkins, H. Davies, Triple junction distribution profiles as assessed by electron backscatter diffraction, Journal of materials science 37 (19) (2002) 4203–4209. doi:10.1023/A:1020052306493.

[61] W. H. Kruskal, Ordinal Measures of Association, Journal of the American Statistical Association 53 (284) (1958) 814–861. doi:10.1080/01621459.1958.10501481.

[62] M. E. Frary, C. A. Schuh, Correlation-space description of the percolation transition in composite microstructures, Physical Review E 76 (4) (2007) 42–45. doi:10.1103/PhysRevE.76.041108.

[63] J. C. Fisher, Calculation of Diffusion Penetration Curves for Surface and Grain Boundary Diffusion, Journal of Applied Physics 22 (1) (1951) 74–77. doi:10.1063/1.1699825.

[64] C. Herring, Surface tension as a motivation for sintering, in: W. E. Kingston (Ed.), The Physics of Powder Metallurgy, McGraw-Hill, New York, 1951, pp. 143–179.

[65] C. E. Rasmussen, C. K. I. Williams, Gaussian Processes for Machine Learning, The MIT Press, Cambridge, MA, 2006.