



**HAL**  
open science

# Energetic upscaling strategy for grain growth. II: Probabilistic macroscopic model identified by Bayesian techniques

Daniel Weisz-Patrault, Sofia Sakout, Alain Ehlacher

► **To cite this version:**

Daniel Weisz-Patrault, Sofia Sakout, Alain Ehlacher. Energetic upscaling strategy for grain growth. II: Probabilistic macroscopic model identified by Bayesian techniques. 2020. hal-03042538v1

**HAL Id: hal-03042538**

**<https://hal.science/hal-03042538v1>**

Preprint submitted on 6 Dec 2020 (v1), last revised 8 Oct 2021 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Energetic upscaling strategy for grain growth. II: Probabilistic macroscopic model identified by Bayesian techniques.

Daniel Weisz-Patrault<sup>a,\*</sup>, Sofia Sakout<sup>a,b</sup>, Alain Ehlacher<sup>b</sup>

<sup>a</sup>*LMS, CNRS, École Polytechnique, Institut Polytechnique de Paris, F-91128 Palaiseau, France*

<sup>b</sup>*Laboratoire Navier, CNRS, École Ponts ParisTech, 6 & 8 Ave Blaise Pascal, F-77455 Marne La Vallée, France*

---

## Abstract

This paper is the second part of an energetic upscaling strategy to simulate grain growth at the macroscopic scale with state variables that contain statistical descriptors of the grain structure. The first part was dedicated to the derivation of a fast mesoscopic model of grain growth based on oriented tessellation updating method, which consists in a succession of Voronoi-Laguerre tessellations obtained by establishing an evolution law directly on the parameters defining the tessellations. In this contribution, the final step of the upscaling strategy is detailed by deriving macroscopic evolutions laws of the state variables representing statistical distributions of the grain structure. The approach relies on macroscopic free energy and dissipation potentials that are identified not axiomatically with experimental calibration, but using a large database of mesoscopic computations. The macroscopic energy is found to be purely deterministic, although the dissipation necessitates to introduce a probabilistic framework. Indeed, an epistemic uncertainty arises due to the loss of information in the reduction of the amount of data between the detailed mesoscopic state and the statistical macroscopic state (i.e., several mesoscopic states can share the same macroscopic state). Classical Bayesian inference has been used to identify the probability density functions associated to the epistemic uncertainty. The resulting stochastic macroscopic model has been compared to several particular mesoscopic evolutions, and good agreement is observed. Thus, this work can be used to simulate grain growth at very large scale with short computation time, while processing rich statistical information about the grain structure, such as mean and standard deviation of the boundary misorientation distribution, grain boundary length density or grain size.

*Keywords:* Grain growth, Upscaling, Macroscopic, Anisotropic grain boundary energy, Bayesian calibration, Stochastic model

---

## 1. Introduction

Grain growth is a thermally activated mechanism that usually occurs after recrystallization during annealing of metals. During grain growth some grains grow at the expense of other grains depending of their respective sizes and crystallographic orientations, which leads to grain coarsening [1]. Classical statistical descriptors of the polycrystalline structure such

---

\*Corresponding author: [daniel.weisz-patrault@polytechnique.edu](mailto:daniel.weisz-patrault@polytechnique.edu)

as morphological and crystallographic textures (e.g., grain size, shape and crystal orientation distributions) evolve during the process. Thus, for different fabrication or forming processes, temperature conditions could be optimized to obtain targeted microstructures, especially for large heterogeneous parts. However, mechanisms involved during grain growth arise at the scale of grain boundaries (GB). Thus, numerical simulations of texture evolution may be difficult for macroscopic parts, which hinders the development of optimization loops to adjust process parameters.

This paper is the second part of an energetic upscaling strategy introduced in [2] and [3]. This strategy aims at developing a model of grain growth at the macroscopic scale that fully relies on finer scales and whose state variables contain statistical descriptors of the grain structure. The proposed upscaling strategy, which involves considering grain growth at various scales, is briefly recalled for the sake of clarity. As shown in figure 1, four typical scales are distinguished: (i) the atomic scale (e.g., crystal lattice and interatomic potential), (ii) the microscopic scale (e.g., grain boundaries), (iii) the mesoscopic scale (e.g., polycrystalline structure) and (iv) the macroscopic scale (statistical descriptors of the grain structure). As energetic concepts are valid at all scales, the upscaling strategy fundamentally relies on various energetic contributions arising at different scales. This energetic upscaling strategy is developed within the framework of standard generalized media [4] that are characterized by their *free energy* and *dissipated power*. These two potentials arise in the energy balance equation combining the first and second laws of thermodynamics, and they depend on macroscopic state variables that should be defined so that the macroscopic state statistically represents the grain structure. The determination of the macroscopic free energy and dissipation potentials as a function of the state variables enables to establish the evolution law of the system at the macroscopic scale. The proposed upscaling strategy consists in determining these two potentials not axiomatically (with parametric functions and calibration with experiments), but on a more physical basis by using a large *database* of computations carried out at the mesoscopic scale. Thus, the macroscopic model emerges from finer scales and is compatible with thermodynamics. It should be

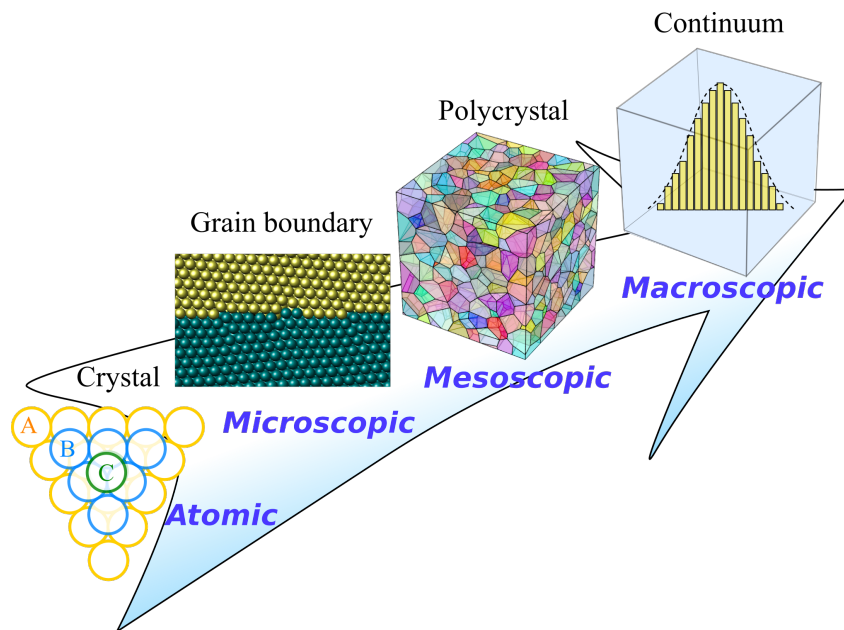


Figure 1: Different scales involved in the upscaling strategy

noted that we do not propose a multiscale approach, for which simulations at the mesoscopic scale are performed online during the macroscopic computation. On the contrary, all the computations at the mesoscopic scale are performed in advance and stored in the database probing the space of polycrystalline structures. On this basis, we can identify the macroscopic free energy and dissipated power as a function of the macroscopic state variables in order to obtain an evolution law that accounts for statistical descriptors of the grain structure.

Therefore, on the one hand the database requires to use intensively a mesoscopic model of grain growth. As a consequence, a sufficiently fast mesoscopic model has been established in the first part of this paper [2]. Many numerical methods have been developed to simulate grain growth at the mesoscopic scale. For instance, models based on cellular automaton and Monte Carlo method [5–9], mobile finite element modeling [10, 11], level set functions [12–14], phase field [15–20], molecular dynamics [21–23], or vertex methods [24–27] have been proposed. In [2] the model relies on updating orientated tessellation method (OTUM), which consists in modeling grain growth by a succession of Voronoi-Laguerre tessellations equipped with a crystal orientation field bu using NEPER [28]. The first contributions in this field are [29, 30].

On the other hand, the final step of the proposed upscaling strategy is broached in the present contribution, and consists in establishing the macroscopic model by using the large database of mesoscopic computations to define macroscopic state variables as well as their macroscopic evolution laws. The main information that is usually upscaled at the macroscopic scale is the average grain size denoted by  $\langle R \rangle$  (i.e., the average equivalent grain radius). Within the context of grain growth, mean field approaches have been continuously developed for several decades in order to establish suitable evolution laws for  $\langle R \rangle$ . Mean field approaches consist in using idealized structures usually with spherical grains representing different grain families in the polycrystalline structure sharing the same properties such as size, GB energy and mobility. However, in real polycrystalline structures, fundamental properties such as misorientation depend not only on the intrinsic characteristics of the grains but also on their connection, that is to say their relative positions. This aspect is neglected in classic mean field approaches, which are based on simplified relationships between the different grain families, as the detailed grain structure is not considered. The early works [31, 32] deriving a growth law for the mean grain radius  $\langle R \rangle$  can be considered as one of the first mean field approaches. This model relies on the curvature driven evolution law, and isotropic GB energy and mobility (i.e., independent on misorientation and GB plane). The classical relationship is obtained:

$$\langle R \rangle \propto t^{\frac{1}{n}} \quad (1)$$

where  $n = 2$  is called the growth exponent. In practice, experiments give evidences that the growth exponent may lie in the following interval [2, 4] [33]. An other famous model called the von Neumann-Mullins (vNM) law [34, 35] relates the number of sides and the area change rate of each grain family. The classical vNM law is formulated within a fully isotropic framework (i.e., isotropic GB energy and isotropic mobility), and relies on the curvature driven evolution law and the assumption that angles at triple junctions are  $120^\circ$ . In addition, an extended vNM law has also been proposed for anisotropic grain growth and tested with a mesoscopic stochastic Monte-Carlo simulations [36]. Among the early mean field theories, the statistical approaches are well known [37]. Considering a grain family of radius  $R$  the following relationship is found

in [37]:

$$\frac{dR}{dt} = m\gamma \left( \frac{1}{\langle R \rangle} - \frac{1}{R} \right) \quad (2)$$

where  $m$  and  $\gamma$  are respectively the constant mobility and GB energy per unit area of the considered grain family, and  $\langle R \rangle$  is the overall average grain radius. This relationship (2) has been obtained within the framework of a more general mean field approach in [38, 39]. By solving (2) Hillert [37] found a steady state statistical distribution of  $R/\langle R \rangle$ . Other well known distributions have been obtained in [40, 41]. At the macroscopic scale (i.e., without referring to mesoscopic quantities) these steady-state distributions are usually the only information. However, a refined analysis based on thermodynamic extremal principle, for which grain families have different GB energies and mobilities, enabled to establish a more general evolution law [42], and the same thermodynamic extremal principle has been used to establish an evolution law directly on the statistic distribution of grain radii [43, 44]. Thus, with an initial distribution, GB energies and mobilities, one can easily determine the evolution of the distribution. These approaches may be considered as a macroscopic model, as grain size distribution evolution is determined without computing a detailed evolution at the mesoscopic scale, even though GB energies and mobilities of each spherical grain are needed. However, since this approach does not take into account the real connectivity between grains (i.e., there is no reference to realistic morphologies), the evolution of boundary misorientation distribution (BMD), GB length density or morphological aspects such as sphericity distributions are not accessible. On the contrary, the approach proposed in this contribution aims at establishing evolution laws for the statistical distribution of the grain structure, including BMDs and GB length density for instance. Semi-topological approaches have also been developed in [45], and enable to introduce some topological and stochastic features in the migration of GBs. However, within the context of grain growth, these semi-topological approaches are still limited to grain size distribution, the main difference with previous approaches being to obtain more realistic distributions. In addition, mean field approaches are usually used to reproduce full field computations (e.g., using level set method or phase field models) to save computation time for industrial processes. A commercial code DIGIMU [46] has been developed, and can perform both full field computations and calibrated mean field estimations.

Thus, the information upscaled at the macroscopic scale is limited to grain size statistics. To overcome this difficulty, the present contribution deals with additional grain statistics, such as mean and standard deviation of misorientations or GB lengths per unit area. The paper is organized as follows. First, the main theoretical results obtained in the mesoscopic model and assumptions are recalled in section 2 for the sake of clarity. The reduction of the amount of data for the construction the macroscopic state variables is broached in section 3. *Primary state variables* are defined to explain GB energy, and an additional *secondary state variable* is introduced to deal with the sensitivity to small grains (i.e., vanishing grains) of a mobility tensor arising in the mesoscopic evolution law. In section 4 the macroscopic evolution laws of both primary and secondary state variables are derived based on the mesoscopic model. These macroscopic evolution laws involve to identify several functions that fully depend on the macroscopic state variables. The identification is performed by using a large database of mesoscopic computations whose content is described in section 5. The function involved in the energy is deterministic and identified in section 6 although the function involved in the mobility is identified in section 7, and is probabilistic. Indeed the reduction of the amount of data leads to an *epistemic uncertainty* that is modeled as a random variable. This epistemic uncertainty

is identified by using Bayesian techniques in section 8. Results are provided in section 9, and additional statistical descriptors of the polycrystalline structure such as the average grain size are estimated as a post-processing. Conclusive remarks are given in section 10.

## 2. Summary of the mesoscopic model

In this section, the main theoretical results and assumptions of the mesoscopic model derived in [2] are briefly recalled for the sake of clarity. The proposed upscaling methodology is applied to plane hexagonal polycrystals. In 3D, this would correspond to face-centered cubic (fcc) crystals, and for each grain the direction [111] is assumed to be aligned with the out of plane direction. Thus, there are three plastic slip systems in 2D. Misorientation between two neighboring grains (characterized by five parameters in 3D) is characterized only by two parameters in 2D: the misorientation angle (denoted by  $\Delta\theta$ ), and the orientation of the GB plane. Thus, the GB energy considered in this paper is computed from fcc crystals sharing the same orientation  $\langle 111 \rangle$  (tilt boundaries) and the effect of the GB plane is neglected. The plane assumption enables to deal with thin structures such as thin films. In addition, since there are 3 plastic slip systems in the plane, the proposed approach also enables to reasonably approximate 3D structures.

Polycrystalline structures are approximated with Voronoi-Laguerre tessellations defined by  $N$  seeds whose dimensionless Cartesian coordinates are denoted by  $(x_j, y_j) \in [0, 1]^2$  and  $N$  dimensionless weights denoted by  $w_j \in \mathbb{R}_+$  (where  $1 \leq j \leq N$ ). In addition,  $N$  crystal orientations  $\theta_j \in [0, \pi/3]$  that represent the  $\langle 111 \rangle$  tilt angles of plane hexagonal lattices (where  $1 \leq j \leq N$ ). The orientated tessellation (OT) is completely determined by the parameter vector  $\alpha = (\mathbf{x}, \mathbf{y}, \mathbf{w}, \boldsymbol{\theta})$  where  $\mathbf{x} = (x_1, \dots, x_N)$ ,  $\mathbf{y} = (y_1, \dots, y_N)$ ,  $\mathbf{w} = (w_1, \dots, w_N)$  and  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_N)$ . Crystal lattice misorientation between grains  $i$  and  $j$  is defined by:

$$\Delta\theta_{ij} = |\theta_j - \theta_i| \in \left[0, \frac{\pi}{3}\right] \quad (3)$$

Furthermore, representative volume elements (RVE) are considered by selecting a subset of  $n$  grains as detailed in [2]. The following evolution law of weights (other parameters being fixed) has been established in [2]:

$$\dot{\mathbf{w}}_{\text{meso}} = -\frac{m(T)}{L_0^3} \underline{\mathbf{M}}_{\text{meso}}(\alpha) \cdot \frac{\partial \mathcal{E}_{\text{meso}}(T, \alpha)}{\partial \mathbf{w}} \quad (4)$$

where  $\mathbf{w}_{\text{meso}}$  are the weights of the OT,  $\mathcal{E}_{\text{meso}}$  is the total energy per unit depth in the RVE,  $m(T)$  is a scalar mobility ( $\text{m}^4 \cdot \text{J}^{-1} \cdot \text{s}^{-1}$ ) that depends on temperature, and  $\underline{\mathbf{M}}_{\text{meso}}$  is a dimensionless mobility second order tensor of size  $n \times n$  depending on the mesoscopic state  $\alpha$ . The index meso refers to the fact that mesoscopic quantities are needed to compute the associated variable. In comparison with [2] the term  $m(T)$  has been factorized although it was included in the second order tensor. It should be noted that  $\underline{\mathbf{M}}_{\text{meso}}$  depends on the choice of the mobility as a function of misorientation at the scale of GBs. In the following the database and then the macroscopic

model are based on the GB energy (5) and the mobility function (6) introduced in [2].

$$\begin{cases} \gamma(0, \Delta\theta) = \gamma_1 \sin\left(\frac{\pi}{2} \frac{\Delta\theta}{\Delta\theta_1}\right) \left[1 - a_1 \ln\left(\sin\left(\frac{\pi}{2} \frac{\Delta\theta}{\Delta\theta_1}\right)\right)\right] \\ (0 \leq \Delta\theta \leq \Delta\theta_1) \\ \gamma(0, \Delta\theta) = \gamma_2 + (\gamma_1 - \gamma_2) \sin\left(\frac{\pi}{2} \frac{\Delta\theta - \frac{\pi}{3}}{\Delta\theta_1 - \frac{\pi}{3}}\right) \left[1 - a_2 \ln\left(\sin\left(\frac{\pi}{2} \frac{\Delta\theta - \frac{\pi}{3}}{\Delta\theta_1 - \frac{\pi}{3}}\right)\right)\right] \\ \left(\Delta\theta_1 \leq \Delta\theta \leq \frac{\pi}{3}\right) \end{cases} \quad (5)$$

where  $\gamma_1 = 0.95 \text{ J.m}^{-2}$ ,  $\gamma_2 = 0.67 \text{ J.m}^{-2}$ ,  $a_1 = 0.5$ ,  $a_2 = 0.1$  and  $\Delta\theta_1 = \pi/6$ .

$$m_{CD}(T, \Delta\theta) = \begin{cases} \frac{m(T)}{X(\Delta\theta)} & \text{if } \min\left\{\Delta\theta, \frac{\pi}{3} - \Delta\theta\right\} \leq \Delta\theta_0 \\ m_{\max}(T) \left(1 - \exp\left[-K \left(\frac{\Delta\theta}{\Delta\theta_1}\right)^p\right]\right) & \text{if } \min\left\{\Delta\theta, \frac{\pi}{3} - \Delta\theta\right\} \geq \Delta\theta_0 \end{cases} \quad (6)$$

where  $\Delta\theta_0 = 4.5\pi/180$  is a threshold between very low angle boundaries and high and intermediate angle boundaries,  $\Delta\theta_1 = \pi/9$ ,  $K = 5$ ,  $p = 4$  and:

$$m(T) = m_{\max}(T) \left(1 - \exp\left[-K \left(\frac{\Delta\theta_0}{\Delta\theta_1}\right)^p\right]\right) X(\Delta\theta_0) \quad (7)$$

and where:

$$X(\Delta\theta) = \frac{6}{\pi} \left(\frac{\pi}{3} + 2\sqrt{3} \ln\left(\frac{\sqrt{3}}{2}\right)\right) \min\left\{\Delta\theta, \frac{\pi}{3} - \Delta\theta\right\} \quad (8)$$

### 3. Macroscopic state variables

#### 3.1. Primary state variables

In this section, the macroscopic state variables considered in this work are introduced and discussed. Since the model identification relies on a database that depends on the choice of the macroscopic state variables, it is convenient to define dimensionless state variables so that the database does not depend on the physical length of the tessellations. All physical quantities are explicitly written as a scaling factor in the dimensionless macroscopic evolution laws. This choice presents the advantage to obtain a dimensionless database (e.g, the same tessellation is not computed several times depending on its physical length) and to demonstrate analytically in the evolution law the influence of physical length, mobility and surface energy. State variables are not defined arbitrarily as the macroscopic energy should be a function of these variables. However, it is a difficult task to determine suitable variables directly from raw data. Thus, intuitive state variables are introduced, and then it is demonstrated that such variables are suitable to describe the macroscopic energy.

Each material point of the macroscopic model should represent a polycrystalline structure (i.e., an RVE) whose energy is the sum of the surface energies carried by the GBs multiplied by their respective length (see [2]). Since the surface energy depends on misorientation, it is expected that the total macroscopic energy at each material point depend on the statistical distributions of GB lengths and misorientations. However, it is clear that complete statistical

distributions constitute a far too rich information to be reasonably processed at each material point of the macroscopic domain. Thus, statistical descriptors (e.g., mean, standard deviation etc.) of length and misorientation distributions should be considered instead of the complete statistical distributions. The *successive statistical moments* of the misorientation distribution weighted by the GB lengths are chosen. Since GB lengths continuously tend to zero when a grain disappear, they are used as weights in the statistical moments so that the statistical moments evolve continuously with respect to time. Indeed, otherwise (i.e., without the GB lengths as weights) the statistical moments would be piecewise constant with a discontinuity each time a grain disappear. Thus, the weighted statistical moments obtained from the mesoscopic state read:

$$\mu_k = \sum_{(i,j) \in I_{\text{RVE}}} l_{ij} \Delta\theta_{ij}^k \quad (9)$$

where  $I_{\text{RVE}}$  is the set of grain boundaries in the RVE,  $l_{ij}$  the dimensionless GB lengths, and  $\Delta\theta_{ij}$  the misorientations. In addition,  $k \in \{0, 1, 2\}$ , indeed as demonstrated in section 6 it is sufficient to consider only the three first statistical moments to accurately account for the total energy.

To give a more intuitive interpretation of  $\mu_k$  ( $k \in \{0, 1, 2\}$ ), one can refer to the length  $L_0$  (m), which has been introduced as the physical side of the OT in [2]. This parameter enables to determine the average grain size. Thus, for instance  $\mu_0/L_0$  is the GB length density (length per unit surface). In addition, the mean of misorientations in the RVE denoted by  $\tilde{\mu}_1$ , and the square of the relative standard deviation<sup>1</sup> denoted by  $\tilde{\mu}_2$  read:

$$\begin{cases} \tilde{\mu}_1 = \frac{\mu_1}{\mu_0} \\ \tilde{\mu}_2 = \frac{\mu_2 \mu_0}{\mu_1^2} - 1 \end{cases} \quad (10)$$

For the sake of clarity, following notations are considered. On the one hand, the statistical moments  $\mu_k$  ( $k \in \{0, 1, 2\}$ ) may be computed from the detailed mesoscopic structure by using (9). In this case, these variables are denoted by  $\boldsymbol{\mu}_{\text{meso}} = (\mu_0, \mu_1, \mu_2)$  and correspond to a reduction of the amount of data with respect to the complete information needed to characterize the OT. Thus,  $\boldsymbol{\mu}_{\text{meso}}$  is a function of  $\boldsymbol{\alpha}$  (where  $\boldsymbol{\alpha} = (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{w}, \boldsymbol{\theta})$  is the mesoscopic state as introduced in [2]). The role of  $\boldsymbol{\mu}_{\text{meso}}$  is twofold: (i) give an overview of mesoscopic evolutions with a limited amount of data, and (ii) provide a tool to define macroscopic state variables within the framework of the upscaling strategy from the mesoscopic scale to the macroscopic scale, as shown in section 4.

On the other hand, the overall aim of this paper is to derive a macroscopic evolution law that do not rely on detailed knowledge of the mesoscopic state. At each material point, only the macroscopic state variables are known, without computing any mesoscopic evolution. The upscaling strategy actually consists in establishing such a macroscopic evolution law. These *primary state variables* are denoted by  $\boldsymbol{\mu}_{\text{macro}}$ , and correspond to a macroscopic version of  $\boldsymbol{\mu}_{\text{meso}}$ , but exist and can be computed independently from the mesoscopic model, that is to say without using the very definition (9). Of course, the identification of the macroscopic model relies on the database of mesoscopic computations, but as soon as this identification is performed, both

---

<sup>1</sup>ratio of the standard deviation over the mean



models are independent. Thus,  $\boldsymbol{\mu}_{\text{macro}}$  does not refer to a specific RVE, but should represent all possible RVEs sharing the same macroscopic state  $\boldsymbol{\mu}_{\text{macro}}$ .

The quality of the proposed upscaling strategy can be evaluated by comparing (i) the overall behavior of detailed mesoscopic evolutions of specific RVEs, obtained by the reduction of the amount of data in (9), and (ii) the corresponding macroscopic evolutions obtained directly from the macroscopic evolution law. Hence the requirement:

$$\begin{cases} \boldsymbol{\mu}_{\text{macro}}(t) \approx \boldsymbol{\mu}_{\text{meso}}(t) \\ \dot{\boldsymbol{\mu}}_{\text{macro}}(t) \approx \dot{\boldsymbol{\mu}}_{\text{meso}}(t) \end{cases} \quad (11)$$

As detailed in the following, the loss of information (due to the fact that several different mesoscopic states can share the same macroscopic state) require to consider the macroscopic evolution law within a probabilistic framework. Thus, the condition (11), which guaranties the upscaling strategy to be reliable should be understood in a probabilistic sense. Intuitively, that is to say that for a particular RVE, the mesoscopic evolution summarized by  $\boldsymbol{\mu}_{\text{meso}}(t)$  should lie in the zone defined by the point-wise dispersion around the mean macroscopic evolution  $\langle \boldsymbol{\mu}_{\text{macro}}(t) \rangle$ , where  $\boldsymbol{\mu}_{\text{macro}}$  is a random variable. Alternatively, considering a large number of different RVEs initially sharing the same overall state  $\boldsymbol{\mu}_{\text{meso}}(t=0)$ , one can compute the empirical mean and standard deviation as a function of time, and compare them to the mean and standard deviation as a function of time of the macroscopic probabilistic evolution.

Consistently with previous notations, for any given quantity denoted by  $Q$ ,  $Q_{\text{meso}}$  refers to the quantity  $Q$  computed from the mesoscopic state  $\boldsymbol{\alpha} = (\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{w}, \boldsymbol{\theta})$ , although  $Q_{\text{macro}}$  refers to the quantity  $Q$  computed from the macroscopic evolution law, without knowledge of the mesoscopic state, hence  $Q_{\text{macro}}$  only depends on macroscopic state variables.

### 3.2. Secondary state variable

In addition, even though it is demonstrated in section 6 that the total energy of a particular RVE is completely determined by  $\boldsymbol{\mu}_{\text{meso}}$ , it is shown in section 7 that the dissipated power requires a probabilistic framework and necessitates the introduction of an additional state variable. Indeed, at the mesoscopic scale the dissipated power depends on a second order mobility tensor  $\underline{\boldsymbol{M}}_{\text{meso}}$  introduced in (4) in section 2. When a grain  $i$  tends to vanish (i.e.  $S_i \rightarrow 0$  where  $S_i$  is the dimensionless surface of the grain), the corresponding grain boundaries vanish (i.e.,  $l_{ij} \rightarrow 0$  with  $j$  denoting the neighboring grains). Since the mobility tensor  $\underline{\boldsymbol{M}}_{\text{meso}}$  includes the inverse of a tensor that linearly depends on the GB lengths, when a grain  $i$  tends to vanish the corresponding components in the mobility tensor  $\underline{\boldsymbol{M}}_{\text{meso}}$  diverge. This could be interpreted as a numerical conditioning issue. However, this behavior actually reflects an instability of small grains that can freely disappear with negligible dissipation cost (i.e., very high mobility). This sensitivity of the mobility tensor to small grains also arises at the macroscopic scale as discussed in section 7. Thus, an additional variable should be added to capture this behavior. To that end, the following variable is introduced:

$$\eta_{\text{meso}} = - \sum_{k=1}^n \log(S_k) \quad (12)$$

where  $S_k$  are the dimensionless grain surfaces in the RVE. At first glance this variable may seem arbitrary. However, it is clear that  $\eta_{\text{meso}}$  rapidly increases when grain sizes decrease.

Unlike  $\mu_{\text{meso}}$ , the variable  $\eta_{\text{meso}}$  is not continuous with respect to time as it diverges each time a grain disappear. This is illustrated in figure 2a, where  $0.5 \eta_{\text{meso}}$  is presented as a function of time with a refined time discretization in order to detect discontinuities (the time step is 10 s). It is clear that  $0.5 \eta_{\text{meso}}$  is fairly well correlated with the evolution of the number of GBs  $n_{GB}$ . This aspect is demonstrated more systematically in figure 2b where  $n_{GB}$  is presented as a function of  $0.5 \eta_{\text{meso}}$  for the tessellations extracted from the database presented in section 5. As

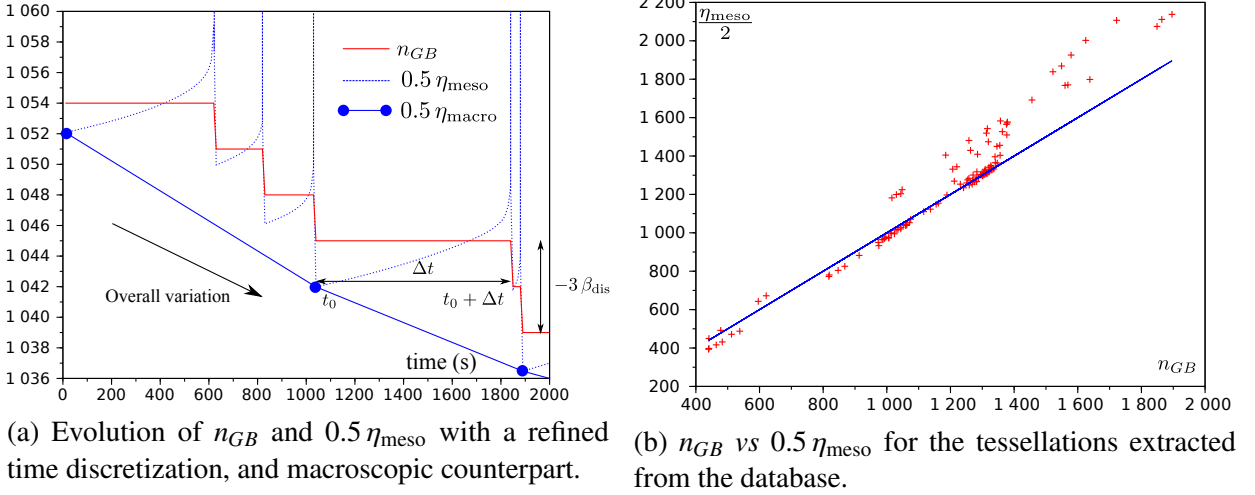


Figure 2: Approximation of  $n_{GB}$  by  $0.5 \eta_{\text{meso}}$ .

for the primary state variables  $\mu_{\text{macro}}$ , a *secondary state variable*  $\eta_{\text{macro}}$  should be determined by establishing a macroscopic evolution law that do not refer to the mesoscopic scale and such as:

$$\eta_{\text{macro}}(t) \approx \eta_{\text{meso}}(t) \quad (13)$$

However, a difficulty arises with the fact that  $\eta_{\text{meso}}$  is discontinuous. Indeed, the time derivative  $\dot{\eta}_{\text{meso}}$  is always positive although the overall variation with respect to time is negative (see. figure 2a), which implies that:

$$\dot{\eta}_{\text{macro}}(t) \neq \dot{\eta}_{\text{meso}}(t) \quad (14)$$

To overcome this difficulty, a smooth approximation of  $\eta_{\text{meso}}(t)$  is sought by considering a larger time scale avoiding the jumps of  $\eta_{\text{meso}}(t)$  as shown in figure 2a. This continuous approximation is obtained by considering the overall variation of  $n_{GB}$  during a finite time interval. More precisely the following idealized situation is considered. During a mesoscopic evolution, consider an instant  $t_0$  belonging to the macroscopic time scale (see figure 2a). In particular small grain sizes are not negligible as  $\eta_{\text{meso}}$  is stable. Consider the grain  $i$ , which is the next grain to disappear. This grain  $i$  is much smaller than the others, and reduces in size. The time needed for the grain  $i$  to disappear is denoted by  $\Delta t$  (see. figure 2a). A crude approximation of  $\Delta t$  is determined by considering a first order Taylor expansion of the dimensionless grain radius  $\sqrt{S_i}/\sqrt{\pi}$ , namely:

$$\sqrt{S_i(t_0 + \Delta t)} \approx \sqrt{S_i(t_0)} + \frac{\dot{S}_i(t_0)}{2\sqrt{S_i(t_0)}} \Delta t = 0 \quad (15)$$

Hence:

$$\Delta t \approx -2 \frac{S_i(t_0)}{\dot{S}_i(t_0)} \quad (16)$$

It should be noted that the choice of  $\sqrt{S_i}$  instead of  $S_i$  to estimate the time needed for the grain to disappear is not arbitrary: the equivalent radius evolves more linearly than the grain surface.

The vanishing grain  $i$  is very likely to be triangular, therefore three GBs disappear. In addition, since  $n_{GB}$  and  $0.5 \eta_{\text{meso}}$  share the same overall evolution, it is imposed that the time derivative of  $0.5 \dot{\eta}_{\text{macro}}$  is equal to the overall variation of  $n_{GB}$  over  $\Delta t$ . However, it can be observed in figure 2a that when a grain  $i$  disappears at  $t_0 + \Delta t$ , several other grains may disappear almost immediately after that. This shorter time scale cannot be considered at the macroscopic scale as a smooth representation of  $n_{GB}$  is sought. To overcome this difficulty, consider  $\beta_{\text{dis}}$  the average number of grains that almost immediately disappear after the grain  $i$ , hence:

$$0.5 \dot{\eta}_{\text{macro}} = \frac{n_{GB}(t_0 + \Delta t) - n_{GB}(t_0)}{\Delta t} = -\frac{3\beta_{\text{dis}}}{\Delta t} \quad (17)$$

At  $t_0$  the only grain that is much smaller than the other is the grain  $i$  (i.e.,  $1/S_i(t_0) \gg S_k(t_0)$  for  $k \neq i$ ), therefore the time derivative of  $\eta_{\text{meso}}$  defined in (12) can be approximated as follows:

$$\dot{\eta}_{\text{meso}}(t_0) = -\sum_{k=1}^n \frac{\dot{S}_k(t_0)}{S_k(t_0)} \approx -\beta_{\text{dis}} \frac{\dot{S}_i(t_0)}{S_i(t_0)} \quad (18)$$

Thus, by using (16), (17) and (18) one obtains:

$$\dot{\eta}_{\text{macro}} \approx -3\beta_{\text{dis}} \dot{\eta}_{\text{meso}}(t_0) \quad (19)$$

Since the previous analysis is very simplified, the factor  $-3\beta_{\text{dis}}$  in (19) is not fully determined as  $\beta_{\text{dis}}$  is unknown. A specific analysis could be carried out to identify  $\beta_{\text{dis}}$  as a function of  $\eta_{\text{meso}}$  (and potentially additional morphological variables). However, this aspect has not been broached in this work. Instead,  $\beta_{\text{dis}}$  is approximated in a phenomenological way, that is to say without rigorous derivation from the mesoscopic scale, as for the rest of the model. The very definition of  $\eta_{\text{meso}}$  (12) shows that  $\beta_{\text{dis}}$  is an increasing function of  $\eta_{\text{meso}}$ . Moreover,  $\beta_{\text{dis}} = 1$  when  $\eta_{\text{meso}}$  is under a certain threshold. Indeed, when a grain  $i$  is disappearing, the minimum number of grains that can disappear is one. These remarks lead to consider a power law with a threshold, and by replacing  $\eta_{\text{meso}}$  by its macroscopic counterpart  $\eta_{\text{macro}}$ :

$$\beta_{\text{dis}} = \begin{cases} \left( \frac{0.5 \eta_{\text{macro}}}{\beta_0} \right)^\delta & 0.5 \eta_{\text{macro}} \geq \beta_0 \\ 1 & 0.5 \eta_{\text{macro}} \leq \beta_0 \end{cases} \quad (20)$$

where  $\delta = 3$  and  $\beta_0 = 1000$  are identified from numerical results presented in section 9 so that  $n_{GB}$  is rather well captured by  $0.5 \eta_{\text{macro}}$ . It should be noted that the other quantities arising in the macroscopic evolution laws derived in section 4 are not identified in a phenomenological way as for  $\beta_{\text{dis}}$ , but in a more theoretical and systematic way as detailed in sections 6 and 7. The simplified phenomenological identification of  $\beta_{\text{dis}}$  is an exception that is proposed to avoid unnecessary technicalities. The choice of  $\beta_{\text{dis}}$  has a significant impact on  $\eta_{\text{macro}}$ , however since  $\mu_{\text{macro}}$  only slightly depends on  $\eta_{\text{macro}}$ , the choice of  $\beta_{\text{dis}}$  has a limited impact on  $\mu_{\text{macro}}$ . Thus, this phenomenological identification seems sufficient.

In addition, the latter relation (19) shows that to overcome the difficulty due to the discontinuity of  $\eta_{\text{meso}}$ , *the database should avoid tessellations with very small grains*, whose sizes are negligible. Indeed, a smooth macroscopic secondary state variable is derived by using  $\dot{\eta}_{\text{meso}}(t_0)$ , which should be finite (i.e., without grains such as  $S_i \approx 0$ ).

## 4. Macroscopic evolution laws

### 4.1. Primary state variables evolution law

In this section, the macroscopic evolution law is derived. Some results related to the database of mesoscopic computations and presented in following sections are required in this section. Nevertheless, the derivation of the macroscopic evolution law is derived first as it enables us to introduce the mathematical quantities to compute in the database.

The mesoscopic evolution law (4), has been derived in [2] from thermodynamical principles. Thus, considering a mesoscopic structure the evolution law (4) is considered. Since  $\boldsymbol{\mu}_{\text{meso}}$  depends on the mesoscopic state the following derivation rule holds:

$$\dot{\boldsymbol{\mu}}_{\text{meso}} = \frac{\partial \boldsymbol{\mu}_{\text{meso}}}{\partial \boldsymbol{w}} \cdot \dot{\boldsymbol{w}}_{\text{meso}} \quad (21)$$

where  $\partial \boldsymbol{\mu}_{\text{meso}} / \partial \boldsymbol{w}$  is a dimensionless second order tensor of size  $3 \times n$ , which can be computed analytically (see. [2]). Hence from (4) and (21):

$$\dot{\boldsymbol{\mu}}_{\text{meso}} = -\frac{m(T)}{L_0^3} \frac{\partial \boldsymbol{\mu}_{\text{meso}}}{\partial \boldsymbol{w}} \cdot \underline{\boldsymbol{M}}_{\text{meso}} \cdot \frac{\partial \mathcal{E}_{\text{meso}}}{\partial \boldsymbol{w}} \quad (22)$$

In addition, in section 6 the analysis of the database of mesoscopic computations shows that the total mesoscopic energy per unit depth denoted by  $\mathcal{E}_{\text{meso}}$  can be very accurately approximated by a deterministic function  $f(\boldsymbol{\mu}_{\text{meso}})$ :

$$\mathcal{E}_{\text{meso}}(T, \boldsymbol{\alpha}) \approx L_0 \gamma_S \frac{G(T)}{G(0)} f(\boldsymbol{\mu}_{\text{meso}}) \quad (23)$$

where  $\gamma_S = 1 \text{ J.m}^{-2}$ ,  $G(T)$  is the temperature dependent shear modulus that captures the temperature dependance of the GB energy, and  $f$  is an analytic function determined in (46) in section 6. Hence, from (4), (22) and (23) one obtains the overall mesoscopic evolution law:

$$\dot{\boldsymbol{\mu}}_{\text{meso}} = -\left( \frac{\gamma_S m(T)}{L_0^2} \frac{G(T)}{G(0)} \right) \left[ \left( \frac{\partial \boldsymbol{\mu}_{\text{meso}}}{\partial \boldsymbol{w}} \right) \cdot \underline{\boldsymbol{M}}_{\text{meso}} \cdot \left( \frac{\partial \boldsymbol{\mu}_{\text{meso}}}{\partial \boldsymbol{w}} \right)^T \right] \cdot \frac{\partial f(\boldsymbol{\mu}_{\text{meso}})}{\partial \boldsymbol{\mu}} \quad (24)$$

On the basis of (24) the following dimensionless second order tensor of size  $3 \times 3$  (called mobility tensor) is introduced:

$$\underline{\boldsymbol{\Gamma}}_{\text{meso}} = \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{w}} \cdot \underline{\boldsymbol{M}}_{\text{meso}} \cdot \left( \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{w}} \right)^T \quad (25)$$

Since the upscaling strategy relies on the objective that the macroscopic state variables  $\boldsymbol{\mu}_{\text{macro}}$  provide a good approximation of the mesoscopic overall state  $\boldsymbol{\mu}_{\text{meso}}$  (see. (11)), the macroscopic energy is defined as the same deterministic function  $f$  but evaluated in  $\boldsymbol{\mu}_{\text{macro}}$ , that is to say:

$$\mathcal{E}_{\text{macro}}(T, \boldsymbol{\mu}_{\text{macro}}) = L_0 \gamma_S \frac{G(T)}{G(0)} f(\boldsymbol{\mu}_{\text{macro}}) \approx \mathcal{E}_{\text{meso}}(T, \boldsymbol{\alpha}) \quad (26)$$

For the energy, the upscaling strategy reduces to the identification of the deterministic dimensionless function  $f$ . In addition, the upscaling strategy also consists in approximating the dimensionless mesoscopic mobility tensor  $\underline{\boldsymbol{\Gamma}}_{\text{meso}}$  by a macroscopic tensor  $\underline{\boldsymbol{\Gamma}}_{\text{macro}}$  defined in (77)

that only depends on the macroscopic state without any knowledge of the detailed mesoscopic state:

$$\underline{\Gamma}_{\text{macro}} \approx \underline{\Gamma}_{\text{meso}} \quad (27)$$

In addition, it is shown in section 7 that  $\underline{\Gamma}_{\text{meso}}$  depends on the secondary variable  $\eta_{\text{meso}}$  defined in (12) and  $\underline{\Gamma}_{\text{macro}}$  requires a probabilistic framework and depends on the secondary state variable  $\eta_{\text{macro}}$ . Thus, by using (11), (26), (27) and (13) into (24), the *macroscopic evolution law of the primary state variables* is obtained:

$$\dot{\boldsymbol{\mu}}_{\text{macro}} = - \left( \frac{\gamma_S m(T) G(T)}{L_0^2 G(0)} \right) \underline{\Gamma}_{\text{macro}}(\boldsymbol{\mu}_{\text{macro}}, \eta_{\text{macro}}) \cdot \frac{\partial f(\boldsymbol{\mu}_{\text{macro}})}{\partial \boldsymbol{\mu}} \quad (28)$$

where  $\partial f(\boldsymbol{\mu}_{\text{macro}})/\partial \boldsymbol{\mu}$  can be interpreted as the dimensionless macroscopic driving force.

#### 4.2. Secondary state variable evolution law

Of course, a macroscopic evolution law is also requested for the secondary state variable  $\eta_M$  in order to update  $\underline{\Gamma}_{\text{macro}}(\boldsymbol{\mu}_{\text{macro}}, \eta_{\text{macro}})$  at each time step. Since grain surfaces  $S_k$  depend on the mesoscopic state, the following derivation rule holds:

$$\dot{\eta}_{\text{meso}} = - \sum_{k=1}^n \frac{\dot{S}_k}{S_k} \quad (29)$$

where:

$$\dot{S}_k = \frac{\partial S_k}{\partial \boldsymbol{w}} \cdot \dot{\boldsymbol{w}}_{\text{meso}} \quad (30)$$

And by using (4) one obtains:

$$\dot{\eta}_{\text{meso}} = \frac{m(T)}{L_0^3} \left( \sum_{k=1}^n \frac{1}{S_k} \frac{\partial S_k}{\partial \boldsymbol{w}} \right) \cdot \underline{\boldsymbol{M}}_{\text{meso}} \cdot \frac{\partial \mathcal{E}_{\text{meso}}}{\partial \boldsymbol{w}} \quad (31)$$

Hence, from (4), (31) and (23) one obtains the overall mesoscopic evolution law:

$$\dot{\eta}_{\text{meso}} = \left( \frac{\gamma_S m(T) G(T)}{L_0^2 G(0)} \right) \left[ \left( \sum_{k=1}^n \frac{1}{S_k} \frac{\partial S_k}{\partial \boldsymbol{w}} \right) \cdot \underline{\boldsymbol{M}}_{\text{meso}} \cdot \left( \frac{\partial \boldsymbol{\mu}_{\text{meso}}}{\partial \boldsymbol{w}} \right)^T \right] \cdot \frac{\partial f(\boldsymbol{\mu}_{\text{meso}})}{\partial \boldsymbol{\mu}} \quad (32)$$

On the basis of (32), the following dimensionless vector of size 3 is introduced:

$$\boldsymbol{\Lambda}_{\text{meso}} = \left( \sum_{k=1}^n \frac{1}{S_k} \frac{\partial S_k}{\partial \boldsymbol{w}} \right) \cdot \underline{\boldsymbol{M}}_{\text{meso}} \cdot \left( \frac{\partial \boldsymbol{\mu}_{\text{meso}}}{\partial \boldsymbol{w}} \right)^T \quad (33)$$

With the same reasoning as before, the upscaling strategy consists in approximating the dimensionless mesoscopic vector  $\boldsymbol{\Lambda}_{\text{meso}}$  by a macroscopic vector  $\boldsymbol{\Lambda}_{\text{macro}}$  defined in (77) that only depends on the macroscopic state without any knowledge of the detailed mesoscopic state:

$$\boldsymbol{\Lambda}_{\text{macro}} \approx \boldsymbol{\Lambda}_{\text{meso}} \quad (34)$$

Thus, the *macroscopic evolution law of the secondary state variable* is obtained:

$$\dot{\eta}_{\text{macro}} = -3\beta_{\text{dis}}(\eta_{\text{macro}}) \left( \frac{\gamma_S m(T) G(T)}{L_0^2 G(0)} \right) \boldsymbol{\Lambda}_{\text{macro}}(\boldsymbol{\mu}_{\text{macro}}, \eta_{\text{macro}}) \cdot \frac{\partial f(\boldsymbol{\mu}_{\text{macro}})}{\partial \boldsymbol{\mu}} \quad (35)$$

where the factor  $-3\beta_{\text{dis}}$  is due to (19).

All terms involved in (28) and (35) are defined from the macroscopic state, and do not refer to the RVE. However,  $L_0$  explicitly arises in (28) and (35), which shows the influence of the physical grain size on grain growth as already discussed in [2]. Of course, it seems that  $L_0$  refers to the RVE, and that state variables could have been defined as densities per unit area instead of being defined as dimensionless quantities in order to discard  $L_0$  from (28) and (35). Nevertheless,  $L_0$  can be seen as a scaling parameter enabling to set the average grain size, which is a macroscopic quantity. In addition to the primary state variables  $\mu_{\text{macro}}$  several other interesting macroscopic quantities such as the average grain size or the average GB length can be computed as a post-processing of the macroscopic evolution law, as shown in section 9.3.

## 5. Database content

In this section, the database of mesoscopic computations is presented. The mesoscopic model based on OTUM has mainly been developed to be intensively used in order to construct a large database that probes the space of OTs. The database is analyzed to identify the energy  $\mathcal{E}_{\text{meso}}$  in (23), the mobility tensor  $\underline{\Gamma}_{\text{meso}}$  in (25), and the vector  $\Lambda_{\text{macro}}$  in (33).

The free software NEPER [28] is used to generate Voronoi-Laguerre tessellations with various morphological statistics. In addition, the probabilistic procedure described in [2] enables us to obtain OTs with various crystallographic textures. In figure 3 three examples of RVEs are presented with grain size and circularity distribution as well as boundary misorientation distribution (BMD).

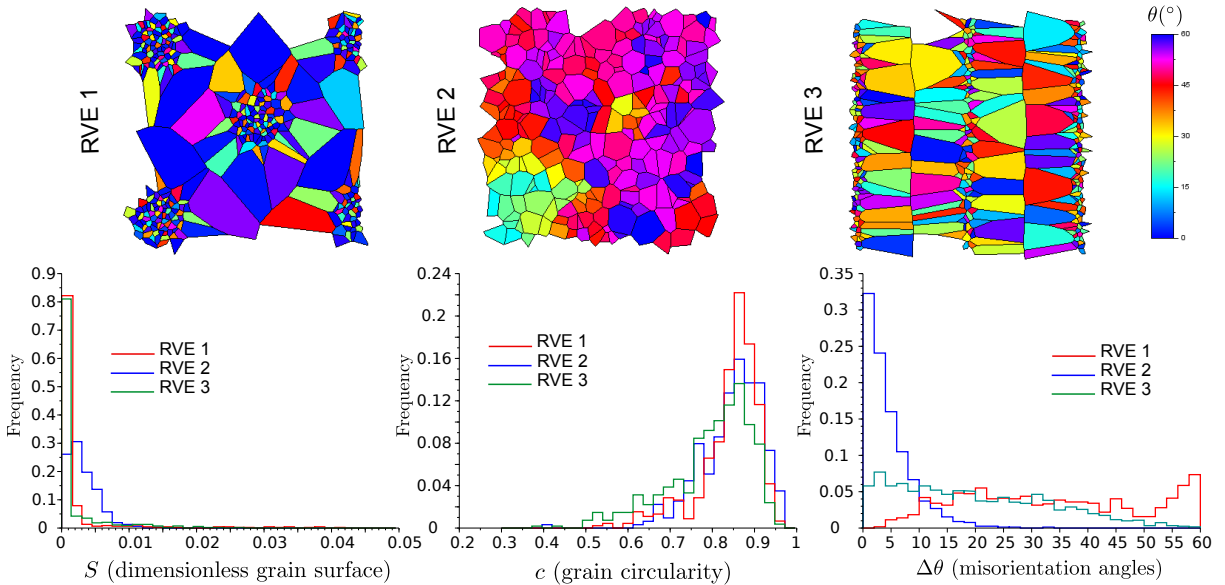


Figure 3: Three different examples of OTs considered for the database, with histograms of grain size, grain circularity, and BMD.

The database that has been used in this work does not include results from mesoscopic evolutions, which would provide successions of OTs that could be included in the database. Instead a more *static* approach has been chosen, the database is constructed from  $N_{\text{tess}} = 121$  different tessellations, each of which is assigned with  $N_{\text{ori}} = 162$  different crystal orientation

fields, leading to a total of  $N_{\text{data}} = 19.602$  OTs. Crystal orientation fields are generated as follows. First, a beta probability density  $\mathcal{B}(\alpha, \beta)$  has been used with  $(\alpha, \beta) \in \{1, \dots, 9\}^2$  (i.e., 81 crystal orientation fields), then a normal probability density  $\mathcal{N}(\mu, \sigma)$  has been used with  $\mu$  and  $\sigma$  describing a homogeneous discretization (with 9 steps) of intervals [15, 45] and [5, 30] respectively (i.e., 81 crystal orientation fields). Of course, several draws of each probability density function involved in the orientation field assignment procedure could be generated instead of only one to enrich the database.

This *static* approach has been chosen to optimize the computation time, as the succession of OTs obtained during a mesoscopic evolution provides rather similar OTs, although the proposed *static* approach enables to probe more efficiently the large space of OTs by selecting very different OTs. But of course, a more *dynamic* approach can also be used to complete the database. When running the mesoscopic code for a given RVE, the intermediate states at each time step can be stored along with the different elements needed for the database, making easy to enrich the database automatically each time the mesoscopic code is used.

The database is constructed on the basis of RVEs obtained from the OTs. Indeed, as discussed in [2], boundary conditions in the mesoscopic model are defined by considering a group of connected grains in the OT that do not belong to the edge of the OT.

Even though the database is dimensionless and does not depend for instance on  $L_0$ , it should be noted that the database obviously depends on the choice of the local GB energy (5) and mobility (6) as a function of misorientation at the scale of GBs. These choices has been discussed in [2].

An extract of the database is given in table 1 where  $i_d$  (with  $1 \leq i_d \leq N_{\text{tess}}$ ) is the index of the Voronoi-Laguerre tessellation,  $k$  (with  $1 \leq k \leq N_{\text{data}}$ ) is the index of each entry in the database, and  $E_{\text{meso}}$  is the dimensionless GB energy at 0 K:

$$E_{\text{meso}} = \frac{\mathcal{E}_{\text{meso}}(T = 0)}{L_0 \gamma_S} \quad (36)$$

The complete database for pure iron that has been constructed in this work is available as a supplementary marterial.

Table 1: Extract of the database of mesoscopic computations for pure iron.

$k$	$i_d$	$n$	$n_{GB}$	$\mu_{\text{meso}}$			$E_{\text{meso}}$	$\Gamma_{\text{meso}}$						$\Lambda_{\text{meso}}$			$\eta_{\text{meso}}$
				$\mu_0$	$\mu_1$	$\mu_2$		$\Gamma_{11}$	$\Gamma_{22}$	$\Gamma_{33}$	$\Gamma_{12}$	$\Gamma_{13}$	$\Gamma_{23}$	$\Lambda_1$	$\Lambda_2$	$\Lambda_3$	
1	1	397	1268	39	8.7	2.4	26.3	1593	387	77	475	153	166	37961	10592	3269	2505
2	1	397	1268	39	8.9	2.7	26.0	1903	521	133	410	116	251	48056	11834	3819	2505
3	1	397	1268	39	9.4	3.1	26.1	1892	550	151	554	190	269	47833	13775	4376	2505
4	1	397	1268	39	10.2	3.8	26.6	1951	969	387	662	312	587	50740	12303	4950	2505
5	1	397	1268	39	9.9	3.7	26.1	1947	1017	349	823	355	561	47579	17698	7287	2505
6	1	397	1268	39	11.1	4.6	27.1	1919	949	420	711	324	595	46292	15144	5529	2505
7	1	397	1268	39	11.5	4.9	27.4	2436	1230	633	996	463	826	60631	20881	8602	2505
8	1	397	1268	39	11.4	5.0	26.9	2445	1371	699	974	427	911	63145	22542	7226	2505
9	1	397	1268	39	11.7	5.3	26.8	2300	1600	911	991	504	1150	66895	36583	20213	2505
10	1	397	1268	39	10.5	3.5	28.7	2686	850	264	806	291	456	66612	19521	6991	2505
11	1	397	1268	39	10.6	3.7	28.1	2198	751	264	585	200	422	55423	14248	4402	2505
12	1	397	1268	39	10.8	4.0	27.8	2430	1104	406	910	381	633	62014	19663	8526	2505
13	1	397	1268	39	11.0	4.4	27.3	2414	1035	445	883	383	640	61547	22430	9771	2505
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

## 6. Identification of the macroscopic free energy

In this section, the macroscopic GB energy per unit depth  $\mathcal{E}_{\text{macro}}$  is identified on the basis of  $\mathcal{E}_{\text{meso}}$  computed for all the OTs listed in the database. The upscaling strategy is conducted in two steps. (i) The first step consists in determining an analytic expression for  $\mathcal{E}_{\text{meso}}$  as a function of  $\boldsymbol{\mu}_{\text{meso}}$  by considering small angle misorientations, which enables to use the analytical local GB energy function proposed by Read & Shockley [47] and a Taylor expansion of the total energy per unit depth in the RVE. (ii) The second step consists in identifying the unknown coefficients involved in the analytic expression exhibited for small angle misorientations (step (i)) by a simple minimization process (i.e., least-squares method) between the energy stored in the database and predictions of the analytic expression. This identification procedure leads to a macroscopic energy function that fit the data almost perfectly, which strongly supports the idea of modeling the macroscopic energy  $\mathcal{E}_{\text{macro}}$  as a deterministic function of  $\boldsymbol{\mu}_{\text{macro}}$ .

### 6.1. Analytic expression for the grain boundary energy

The total mesoscopic GB energy defined in [2] is recalled here for the sake of clarity:

$$\mathcal{E}_{\text{meso}}(T, \boldsymbol{\alpha}) = L_0 \frac{G(T)}{G(0)} \sum_{(i,j) \in I_{\text{RVE}}} l_{ij} \gamma(0, \Delta\theta_{ij}) \quad (37)$$

where  $\gamma(0, \Delta\theta_{ij})$  is the local anisotropic GB energy per unit area at 0 K. If small angle misorientations are considered the Read & Shockley formula [47] is accurate, and presents the advantage to deal with simple analytic expressions, and then:

$$\gamma(0, \Delta\theta) = \gamma_S \kappa_1 \Delta\theta (\kappa_2 - \log(\Delta\theta)) \quad (38)$$

where  $\gamma_S = 1 \text{ J.m}^{-2}$ , and  $\kappa_1, \kappa_2$  are dimensionless parameters. Hence:

$$\mathcal{E}_{\text{meso}}(T, \boldsymbol{\alpha}) = L_0 \frac{G(T)}{G(0)} \gamma_S \kappa_1 \sum_{(i,j) \in I_{\text{RVE}}} l_{ij} \Delta\theta_{ij} (\kappa_2 - \log(\Delta\theta_{ij})) \quad (39)$$

Moreover,  $\epsilon_{ij}$  denotes the deviation of the misorientation  $\Delta\theta_{ij}$  from the mean misorientation  $\tilde{\mu}_1$  defined in (10), hence:

$$\Delta\theta_{ij} = \tilde{\mu}_1 + \epsilon_{ij} \quad (40)$$

Thus, combining (39) and (40) one obtains:

$$\mathcal{E}_{\text{meso}}(T, \boldsymbol{\alpha}) = L_0 \frac{G(T)}{G(0)} \gamma_S \kappa_1 \sum_{(i,j) \in I_{\text{RVE}}} l_{ij} (\tilde{\mu}_1 + \epsilon_{ij}) \left( \kappa_2 - \log(\tilde{\mu}_1) - \log\left(1 + \frac{\epsilon_{ij}}{\tilde{\mu}_1}\right) \right) \quad (41)$$

By considering that the deviations  $\epsilon_{ij}$  are small compared to  $\tilde{\mu}_1$  (i.e.,  $\epsilon_{ij}/\tilde{\mu}_1 \ll 1$ ), the Taylor expansion at the second order in  $\epsilon_{ij}$  reads:

$$\mathcal{E}_{\text{meso}}(T, \boldsymbol{\alpha}) = L_0 \frac{G(T)}{G(0)} \gamma_S \kappa_1 \sum_{(i,j) \in I_{\text{RVE}}} l_{ij} (\tilde{\mu}_1 + \epsilon_{ij}) \left( \kappa_2 - \log(\tilde{\mu}_1) - \frac{\epsilon_{ij}}{\tilde{\mu}_1} + \frac{\epsilon_{ij}^2}{2\tilde{\mu}_1^2} + o(\epsilon_{ij}^2) \right) \quad (42)$$

Moreover, using the definition (40), the following equations can be easily verified:

$$\begin{cases} \sum_{(i,j) \in I_{\text{RVE}}} l_{ij} \epsilon_{ij} = 0 \\ \sum_{(i,j) \in I_{\text{RVE}}} l_{ij} \epsilon_{ij}^2 = \mu_2 - \frac{\mu_1^2}{\mu_0} \end{cases} \quad (43)$$



Hence:

$$\mathcal{E}_{\text{meso}}(T, \alpha) \approx L_0 \frac{G(T)}{G(0)} \gamma_S \kappa_1 \mu_0 \tilde{\mu}_1 \left( \kappa_2 - \log(\tilde{\mu}_1) - \frac{\mu_2 \mu_0}{2\mu_1^2} \right) \quad (44)$$

By using the definition of the squared relative standard deviation introduced in (10), the total GB energy for small angle misorientations can be approximated as follows:

$$\mathcal{E}_{\text{meso}}(T, \alpha) \approx L_0 \frac{G(T)}{G(0)} \gamma_S f(\boldsymbol{\mu}_{\text{meso}}) \quad (45)$$

where the dimensionless function  $f$  reads:

$$f(\mu_0, \mu_1, \mu_2) = \kappa_1 \mu_0 \tilde{\mu}_1 \left( \kappa_2 - \log(\tilde{\mu}_1) - \frac{1}{2} (\tilde{\mu}_2 + 1) \right) \quad (46)$$

The simple analytic expression (45) is used to fit the energy stored in the database as a function of  $\boldsymbol{\mu}_{\text{meso}}$ . The two coefficients  $\kappa_1$  and  $\kappa_2$  need to be identified.

## 6.2. Identification of the unknown coefficients

As introduced in [2], the local GB energy used in the model is not the Read & Shockley formula [47] but an analytic function (5) proposed by Wolf [48], which enabled to interpolate molecular dynamics computations. However, for low angle misorientations the interpolation function matches the Read & Shockley formula and the analysis presented in section 6.1 remains valid. In addition, even though the analysis has been conducted for very small angle misorientations, the obtained analytic function  $f$  in (46) is used as an educated guess to interpolate the energy in the database even for large angle misorientations.

In the following, a minimization is performed to identify  $\kappa_1$  and  $\kappa_2$ , in order to fully determine the function  $f$  defined in (46). This function is essential to define the macroscopic energy  $\mathcal{E}_{\text{macro}}$  as detailed in (26). The data set used for the minimization is composed of the following components of the database:

$$U = \left\{ (\mu_0^{(k)}, \mu_1^{(k)}, \mu_2^{(k)}, E_{\text{meso}}^{(k)}), 1 \leq k \leq N_{\text{data}} \right\} \quad (47)$$

where  $E_{\text{meso}}$  is the dimensionless GB energy at 0 K defined in (36), and the exponent ( $k$ ) stands for the  $k$ -th entry in the database. The least-squares method is used to carry out the minimization procedure. More precisely, the natural Euclidean distance between the dimensionless GB energy  $E_{\text{meso}}$  in the database and predictions of the analytical function  $f$  in (46) is minimized. Thus the minimization problem reads:

$$(\kappa_1, \kappa_2) = \underset{(\kappa_1^*, \kappa_2^*) \in \mathbb{R}^2}{\operatorname{argmin}} \sum_{k=1}^{N_{\text{data}}} \left( E_{\text{meso}}^{(k)} - \mathcal{F}^{(k)}[\kappa_1^*, \kappa_2^*] \right)^2 \quad (48)$$

where:

$$\mathcal{F}^{(k)}[\kappa_1^*, \kappa_2^*] = \kappa_1^* \mu_0^{(k)} \tilde{\mu}_1^{(k)} \left( \kappa_2^* - \log(\tilde{\mu}_1^{(k)}) - \frac{1}{2} (\tilde{\mu}_2^{(k)} + 1) \right) \quad (49)$$

where  $\tilde{\mu}_1^{(k)}, \tilde{\mu}_2^{(k)}$  are computed from  $\mu_0^{(k)}, \mu_1^{(k)}, \mu_2^{(k)}$  accordingly to (10).

The linear minimization problem (48) is fairly simple and the solution reads:

$$\begin{bmatrix} \kappa_1 & \kappa_2 \\ \kappa_1 & \kappa_2 \end{bmatrix} = (\underline{\mathbf{X}}^T \cdot \underline{\mathbf{X}})^{-1} \cdot \underline{\mathbf{X}}^T \cdot \mathbf{E}_{\text{meso}} \quad (50)$$

where  $\mathbf{E}_{\text{meso}} = (E_{\text{meso}}^{(k)})_{1 \leq k \leq N_{\text{data}}}$  is a vector of size  $N_{\text{data}}$ , and  $\underline{\mathbf{X}}$  is a second order tensor of size  $N_{\text{data}} \times 2$  defined by:

$$\begin{cases} X_{k,1} = \mu_0^{(k)} \tilde{\mu}_1^{(k)} \\ X_{k,2} = -\mu_0^{(k)} \tilde{\mu}_1^{(k)} \left( \log(\tilde{\mu}_1^{(k)}) + \frac{1}{2} (\tilde{\mu}_2^{(k)} + 1) \right) \end{cases} \quad (51)$$

Results are listed in table 2.

Table 2: Dimensionless coefficients involved in the energy.

$\kappa_1$	1.7
$\kappa_2$	0.9

The coefficient of determination  $r^2$  quantifies the part of the variance of the data (i.e.,  $E_{\text{meso}}$ ) that is explained by the model (i.e., the  $f$  function). This coefficient of determination has been computed, and  $r^2 = 0.998$ . Thus, the dimensionless analytic function  $f$  almost perfectly captures the GB energy at the macroscopic scale, which can be considered as deterministic function. A comparison between data and the model is presented in figure 4, and a perfect fit is observed.

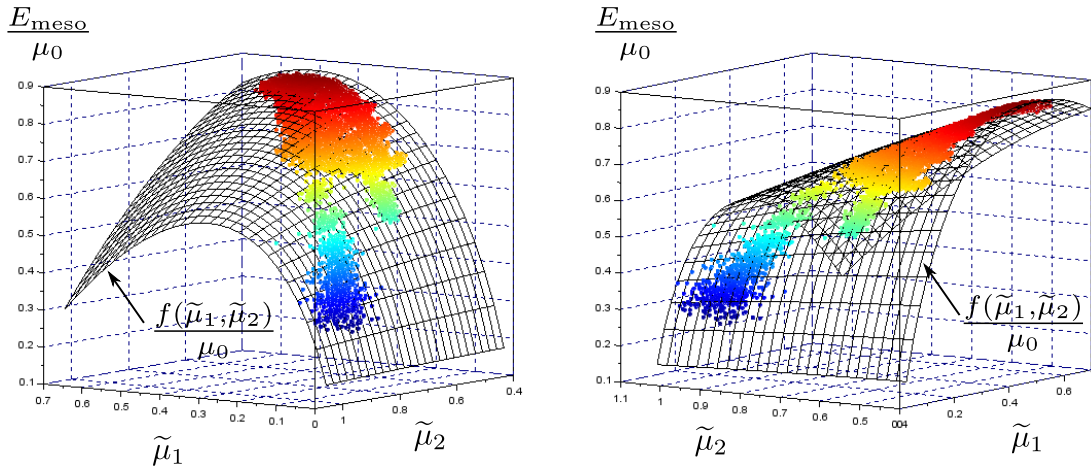


Figure 4: Dimensionless GB energy  $E_{\text{meso}}/\mu_0$  and analytic function  $f/\mu_0$  as a function of  $\tilde{\mu}_1$  and  $\tilde{\mu}_2$  computed from  $\boldsymbol{\mu}_{\text{meso}}$ .

The main result of this section is that the upscaling of the total GB energy leads to a deterministic function of the macroscopic state. Therefore, the choice of considering the successive statistical moments in (9) as a basis to construct the primary state variables  $\boldsymbol{\mu}_{\text{macro}}$  is validated.

## 7. Identification of the macroscopic mobility tensor

In this section, the mobility tensor  $\underline{\boldsymbol{\Gamma}}_{\text{meso}}$  arising in(25) and the vector  $\boldsymbol{\Lambda}_{\text{meso}}$  arising in (33) are analyzed. There is no available analytical computation such as presented in section 6.1 to guide the development of a macroscopic model. Thus, the identification is carried out only with statistical treatments of the data. In addition, as already mentioned in section 3.2, a secondary

state variable  $\eta_{\text{macro}}$  is necessary to deal with the sensitivity of  $\underline{\Gamma}_{\text{meso}}$  and  $\underline{\Lambda}_{\text{meso}}$  to the presence of very small grains. The secondary state variable is introduced to explain a new quantity that will be called the *tessellation amplification factor*, which is a scaling factor so that all tessellations are comparable.

### 7.1. Raw data

In figure 5, raw data of the dimensionless mobility tensor  $\underline{\Gamma}_{\text{meso}}$  and the vector  $\underline{\Lambda}_{\text{meso}}$  are presented as a function of  $\tilde{\mu}_1$  (computed from  $\mu_{\text{meso}}$  according to (10)). Colors correspond to the tessellation index. More precisely, each point in figure 5 correspond to a specific OT indexed by  $k$  (with  $1 \leq k \leq N_{\text{data}}$ ) and all points sharing the same color correspond to a single tessellation indexed by  $i_d$  (where  $1 \leq i_d \leq N_{\text{tess}}$ ) with different BMDs. It should be noted that the variable  $\tilde{\mu}_2$  that has been used for the energy (see. figure 4) does not explain the variance of the raw data for  $\underline{\Gamma}_{\text{meso}}$  and  $\underline{\Lambda}_{\text{meso}}$ , and is not used to interpret these results. In addition, it is

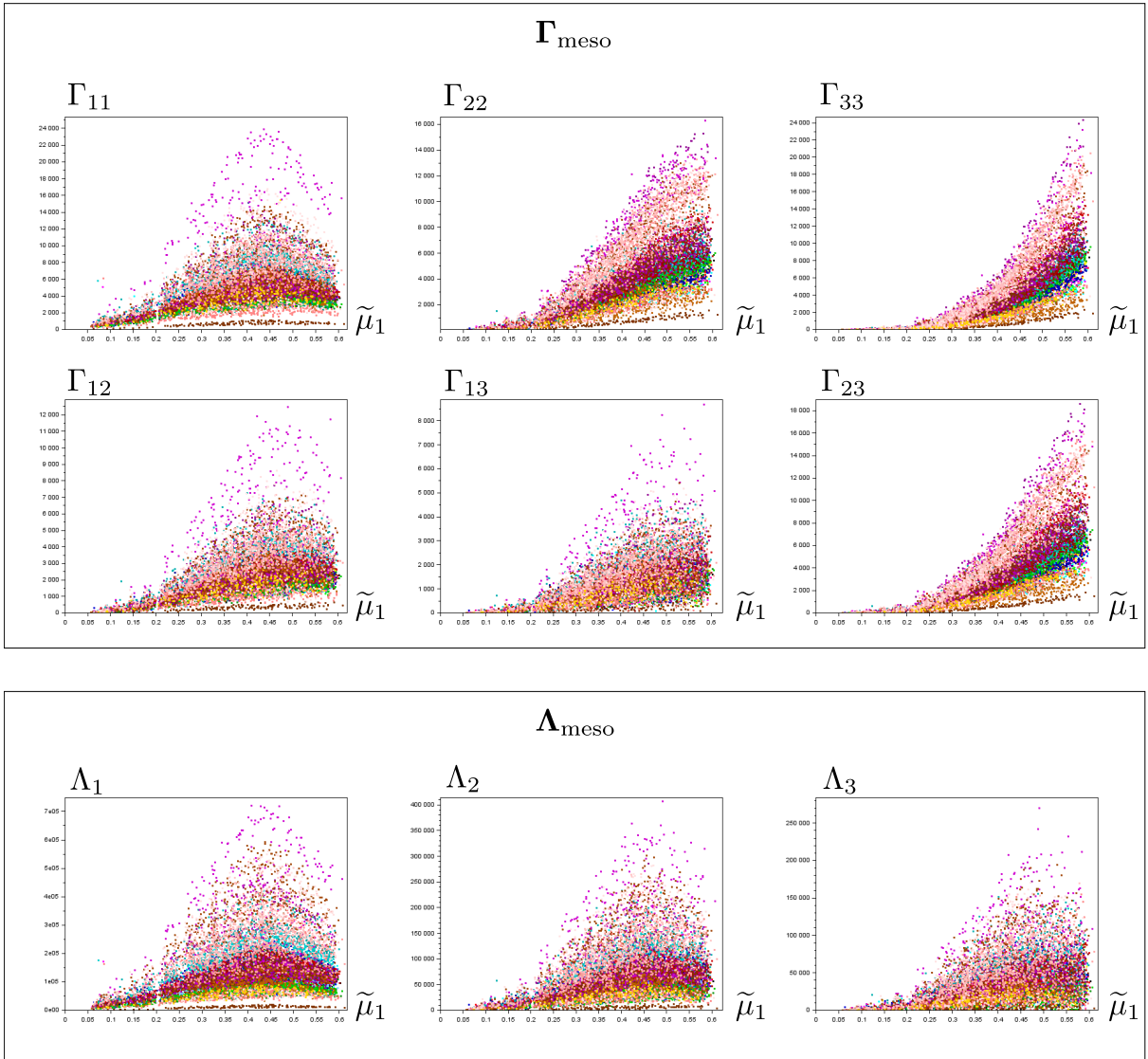


Figure 5: Raw data of the dimensionless mobility tensor  $\underline{\Gamma}_{\text{meso}}$  and the vector  $\underline{\Lambda}_{\text{meso}}$  as a function of  $\tilde{\mu}_1$  computed from  $\mu_{\text{meso}}$ . Colors correspond to the tessellation index  $i_d$  ( $1 \leq i_d \leq N_{\text{tess}}$ ).

clear from figure 5 that the dispersion of the raw data is very significant, which would hinder the development of the macroscopic mobility tensor  $\underline{\Gamma}_{\text{macro}}$  and vector  $\underline{\Lambda}_{\text{macro}}$  involved in the macroscopic evolution laws (28) and (35) respectively. However, it seems that the tessellation index  $i_d$  explains a significant part of the variance of the raw data, thus there is a quantity that only depends on the tessellation (and does not depend on BMD) that explains a large part of the dispersion observed in figure 5, this quantity is introduced as a scaling factor in the following. This aspect is evidenced in logarithmic scales as shown in figure 6. Indeed, for each component the behaviors corresponding to the different tessellations  $i_d$  are very similar for all tessellations up to an offset in logarithmic scales, which shows that a *tessellation amplification factor* can be introduced in the usual scales in order to reduce data dispersion.

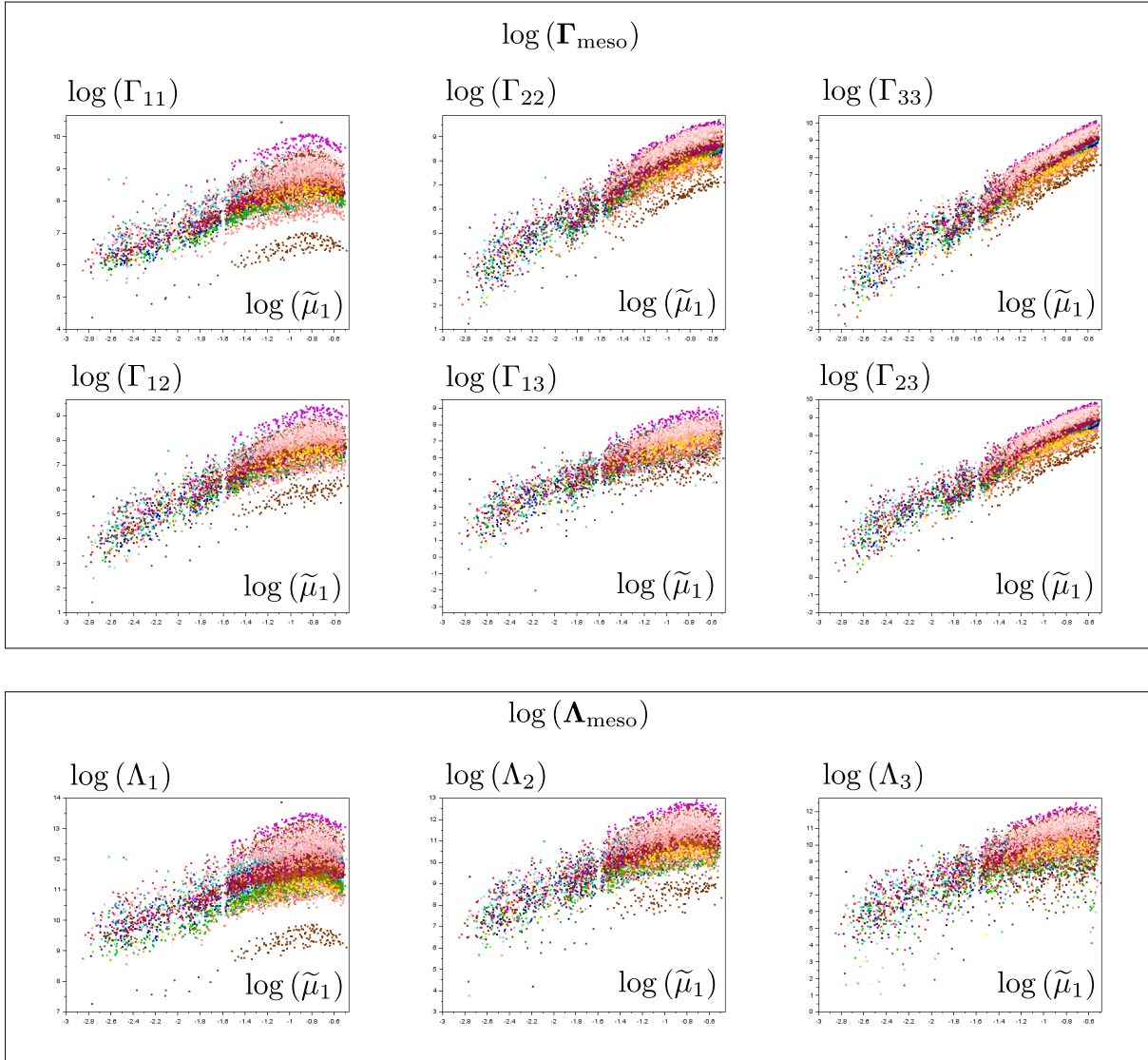


Figure 6: Raw data of the dimensionless mobility tensor  $\underline{\Gamma}_{\text{meso}}$  and vector  $\underline{\Lambda}_{\text{meso}}$  as a function of  $\tilde{\mu}_1$  computed from  $\underline{\mu}_{\text{meso}}$  in logarithmic scales. Colors correspond to the tessellation index  $i_d$  ( $1 \leq i_d \leq N_{\text{tess}}$ ).

## 7.2. Interpolation in logarithmic scales and amplification factor

The raw data in logarithmic scales are almost linear as a function of  $\log \tilde{\mu}_1$ , as shown in figure 6. More precisely, it is convenient to interpolate these data with two linear functions that

smoothly merge at  $\log(\tilde{\mu}_1) = x_{\text{cut}}$  where  $x_{\text{cut}} = -0.82$ . Thus, sigmoid functions are considered for the interpolation to carry out the smooth connection between the two linear functions. Thus the interpolation function in logarithmic scales, for each tessellation  $i_d$ , and each component reads:

$$F_{(\cdot)}^{(i_d)} : x \mapsto \frac{A_{(\cdot)}^{(i_d)} x + B_{(\cdot)}^{(i_d)}}{1 + \exp(a(x - x_{\text{cut}}))} + \frac{C_{(\cdot)}^{(i_d)} x + D_{(\cdot)}^{(i_d)}}{1 + \exp(-a(x - x_{\text{cut}}))} \quad (52)$$

where the symbol  $(\cdot)$  stands for the index of the component (i.e., 11,22, etc. for  $\underline{\Gamma}_{\text{meso}}$  and 1,2,3 for  $\underline{\Lambda}_{\text{meso}}$ ), and  $a$  is a parameter that determines the behavior of the connection between the two linear functions  $x \mapsto A_{(\cdot)}^{(i_d)} x + B_{(\cdot)}^{(i_d)}$  and  $x \mapsto C_{(\cdot)}^{(i_d)} x + D_{(\cdot)}^{(i_d)}$ . This parameters has been set to  $a = 10$ . Coefficients  $A_{(\cdot)}^{(i_d)}, B_{(\cdot)}^{(i_d)}, C_{(\cdot)}^{(i_d)}, D_{(\cdot)}^{(i_d)}$  are not independent as the linear functions should connect in  $x = x_{\text{cut}}$ . Thus the following condition holds:

$$D_{(\cdot)}^{(i_d)} = (A_{(\cdot)}^{(i_d)} - C_{(\cdot)}^{(i_d)}) x_{\text{cut}} + B_{(\cdot)}^{(i_d)} \quad (53)$$

Thus, coefficients  $A_{(\cdot)}^{(i_d)}, B_{(\cdot)}^{(i_d)}, C_{(\cdot)}^{(i_d)}$  should be determined by minimizing a distance between the data and the predictions of the interpolation functions  $F_{(\cdot)}^{(i_d)}$  defined in (52). The least-squares method is used to carry out the minimization, thus:

$$(A_{(\cdot)}^{(i_d)}, B_{(\cdot)}^{(i_d)}, C_{(\cdot)}^{(i_d)}) = \underset{(A^*, B^*, C^*) \in \mathbb{R}^4}{\text{argmin}} \sum_{k \in \mathcal{K}(i_d)} [\log(\Gamma_{(\cdot)}^{(k)}) - F_{(\cdot)}^{(i_d)}(\log(\tilde{\mu}_1^{(k)}))]^2 \quad (54)$$

where  $\Gamma_{(\cdot)}^{(k)}$  is the component  $(\cdot)$  of  $\underline{\Gamma}_{\text{meso}}$  for the  $k$ -th OT in the database, and where  $\mathcal{K}(i_d)$  is the set of OTs sharing the tessellation  $i_d$ . Thus,  $\text{card}[\mathcal{K}(i_d)] = N_{\text{ori}}$ , where  $N_{\text{ori}}$  is the number of crystallographic orientation fields for each tessellation  $i_d$ . Of course the same minimization problem holds for the components of  $\underline{\Lambda}_{\text{meso}}$ . The minimization problem (54) is linear and the solution reads:

$$\begin{bmatrix} A_{(\cdot)}^{(i_d)} \\ B_{(\cdot)}^{(i_d)} \\ C_{(\cdot)}^{(i_d)} \end{bmatrix} = (\underline{\mathbf{X}}^T \cdot \underline{\mathbf{X}})^{-1} \cdot \underline{\mathbf{X}}^T \cdot \mathbf{\Gamma}_{(\cdot)}^{(i_d)} \quad (55)$$

where  $\mathbf{\Gamma}_{(\cdot)}^{(i_d)} = (\Gamma_{(\cdot)}^{(k)})_{k \in \mathcal{K}(i_d)}$ . The vector  $\mathbf{\Gamma}_{(\cdot)}^{(i_d)}$  is of size  $N_{\text{ori}}$ , and  $\underline{\mathbf{X}}$  is a second order tensor of size  $N_{\text{ori}} \times 3$  defined by:

$$\forall k \in \mathcal{K}(i_d), \begin{cases} X_{k,1} = \frac{\log(\tilde{\mu}_1^{(k)})}{(1 + \exp(a(\log(\tilde{\mu}_1^{(k)}) - x_{\text{cut}})))} + \frac{x_{\text{cut}}}{(1 + \exp(-a(\log(\tilde{\mu}_1^{(k)}) - x_{\text{cut}})))} \\ X_{k,2} = \frac{1}{(1 + \exp(a(\log(\tilde{\mu}_1^{(k)}) - x_{\text{cut}})))} + \frac{1}{(1 + \exp(-a(\log(\tilde{\mu}_1^{(k)}) - x_{\text{cut}})))} \\ X_{k,3} = \frac{(\log(\tilde{\mu}_1^{(k)}) - x_{\text{cut}})}{(1 + \exp(-a(\log(\tilde{\mu}_1^{(k)}) - x_{\text{cut}})))} \end{cases} \quad (56)$$

Of course the same minimization procedure holds for the components of  $\underline{\Lambda}_{\text{meso}}$ . For instance, in figure 7a the interpolation is demonstrated for some tessellations for the 11 component of  $\underline{\Gamma}_{\text{meso}}$ . It is clear that all tessellations have similar interpolation up to an offset. Thus in classic scales there is a *tessellation amplification factor* denoted by  $\zeta_{(\cdot)}$ , which leads in logarithmic scales to

an additive offset  $\log(\zeta_{(\cdot)})$ . This quantity is specific to each tessellation and does not depend on the crystal orientation distribution in the RVE. Thus, in the following the notation  $\zeta_{(\cdot)}^{(i_d)}$  refers to the specific value of  $\zeta_{(\cdot)}$  for the tessellation  $i_d$  where  $1 \leq i_d \leq N_{\text{tess}}$ . For the first tessellation (i.e.,  $i_d = 1$ ) the offset is arbitrarily set to:

$$\log(\zeta_{(\cdot)}^{(1)}) = F_{(\cdot)}^{(1)}(x_{\text{cut}}) \quad (57)$$

For the other tessellations (i.e.,  $i_d \geq 2$ ), the offset  $\log(\zeta_{(\cdot)}^{(i_d)})$  is determined by solving the following least-squares minimization problem:

$$\log(\zeta_{(\cdot)}^{(i_d)}) = \underset{\log(\zeta^*) \in \mathbb{R}}{\operatorname{argmin}} \sum_{k \in \mathcal{K}(i_d)} \left[ F_{(\cdot)}^{(i_d)}(\log(\tilde{\mu}_1^{(k)})) - \log(\zeta^*) - (F_{(\cdot)}^{(1)}(\log(\tilde{\mu}_1^{(k)})) - \log(\zeta_{(\cdot)}^{(1)})) \right]^2 \quad (58)$$

The minimization problem (58) is linear, and the solution reads for  $i_d \geq 2$ :

$$\log(\zeta_{(\cdot)}^{(i_d)}) = \frac{1}{N_{\text{ori}}} \sum_{k \in \mathcal{K}(i_d)} F_{(\cdot)}^{(i_d)}(\log(\tilde{\mu}_1^{(k)})) - (F_{(\cdot)}^{(1)}(\log(\tilde{\mu}_1^{(k)})) - \log(\zeta_{(\cdot)}^{(1)})) \quad (59)$$

Of course (59) also holds for the components of  $\mathbf{\Lambda}_{\text{meso}}$ . For instance, in figure 7b the interpolation corrected with the tessellation amplification factor is demonstrated for the 11 component of  $\mathbf{\Gamma}_{\text{meso}}$ .

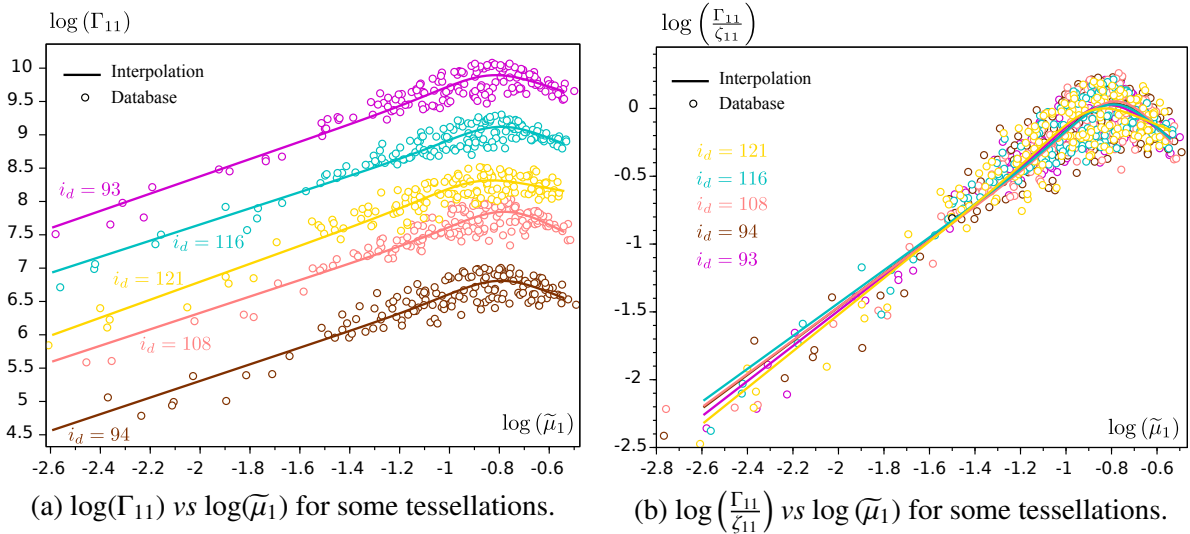


Figure 7: Comparison between data and interpolation for some tessellations.

### 7.3. Interpolation of rescaled data

The rescaled data are by definition:

$$\mathcal{S} = \left\{ \left( \frac{\Gamma_{11}^{(k)}}{\zeta_{11}^{(i_d)}}, \frac{\Gamma_{22}^{(k)}}{\zeta_{22}^{(i_d)}}, \frac{\Gamma_{33}^{(k)}}{\zeta_{33}^{(i_d)}}, \frac{\Gamma_{12}^{(k)}}{\zeta_{12}^{(i_d)}}, \frac{\Gamma_{13}^{(k)}}{\zeta_{13}^{(i_d)}}, \frac{\Gamma_{23}^{(k)}}{\zeta_{23}^{(i_d)}}, \frac{\Lambda_1^{(k)}}{\zeta_{11}^{(i_d)}}, \frac{\Lambda_2^{(k)}}{\zeta_{22}^{(i_d)}}, \frac{\Lambda_3^{(k)}}{\zeta_{33}^{(i_d)}} \right), \forall k \in \mathcal{K}(i_d), \forall i_d \in \{1, \dots, N_{\text{tess}}\} \right\} \quad (60)$$

Since the rescaled data coincide on a common trend (see. figure 7b), it is possible to propose a single interpolation function in logarithmic scales for all the tessellations:

$$F_{(\cdot)}^{\text{sc}} : x \mapsto \frac{A_{(\cdot)}x + B_{(\cdot)}}{1 + \exp(a(x - x_{\text{cut}}))} + \frac{C_{(\cdot)}x + D_{(\cdot)}}{1 + \exp(-a(x - x_{\text{cut}}))} \quad (61)$$

where the exponent sc stands for *re-scaled*, and the symbol  $(\cdot)$  stands for 11, 22, etc. for the components of  $\underline{\Gamma}_{\text{meso}}$  and for 1,2,3 for the components of  $\underline{\Lambda}_{\text{meso}}$  and where:

$$D_{(\cdot)} = (A_{(\cdot)} - C_{(\cdot)}) x_{\text{cut}} + B_{(\cdot)} \quad (62)$$

Coefficients  $A_{(\cdot)}, B_{(\cdot)}, C_{(\cdot)}$  are determined in a very similar way as in section 7.2. The first difference is that the rescaled data (60) are considered instead of the raw data. The second difference is that for each component of  $\underline{\Gamma}_{\text{meso}}$  and  $\underline{\Lambda}_{\text{meso}}$  instead of identifying an interpolation for each tessellation, a single interpolation is identified for all the tessellations considered as a whole. The rescaled data (60) are presented along with the interpolation (61) in logarithmic scales in figure 8 and in classic scales in figure 9. The identified coefficients are listed in table 3.

Table 3: Coefficients for the interpolation of rescaled data.

Component	$A_{(\cdot)}$	$B_{(\cdot)}$	$C_{(\cdot)}$
11	1.238	1.031	-0.926
22	2.585	2.245	1.129
33	3.991	3.534	3.459
12	1.925	1.609	-0.402
13	2.506	2.086	0.238
23	3.296	2.898	2.332
1	1.244	1.043	-0.986
2	1.944	1.643	-0.58
3	2.554	2.174	-0.239

Moreover, it should be noted that the different components of  $\zeta_{(\cdot)}$  are strongly correlated. In figure 10 linear relationships are exhibited. Thus, in order to simplify the subsequent analysis, only  $\zeta_{33}$  is analyzed in the following, as the other components can be inferred with reasonable accuracy from  $\zeta_{33}$ . Proportionality coefficients are listed in table 4. Of course, this analysis is simplified and a multivariate model with a covariance could have been used instead, nevertheless the correlation between the different components of the tessellation amplification factor  $\zeta_{(\cdot)}$  is sufficient to enable such an assumption.

#### 7.4. Macroscopic amplification factor

At this stage, the interpolation (61) can be used at the macroscopic scale. However, the tessellation amplification factor  $\zeta_{(\cdot)}$  obtained in (59) highly depends on the detailed arrangement of the tessellation. An interpolation function that can be used at the macroscopic scale is derived in this section. This interpolation relies on the variable  $\eta_{\text{meso}}$  introduced in (12). In figure 11 the values of  $\zeta_{33}$  obtained for all tessellations in (59) are presented as a function of  $\eta_{\text{meso}}$ . A correlation can be observed and a linear interpolation function is proposed:

$$G : x \mapsto \widehat{A}x + \widehat{B} \quad (63)$$

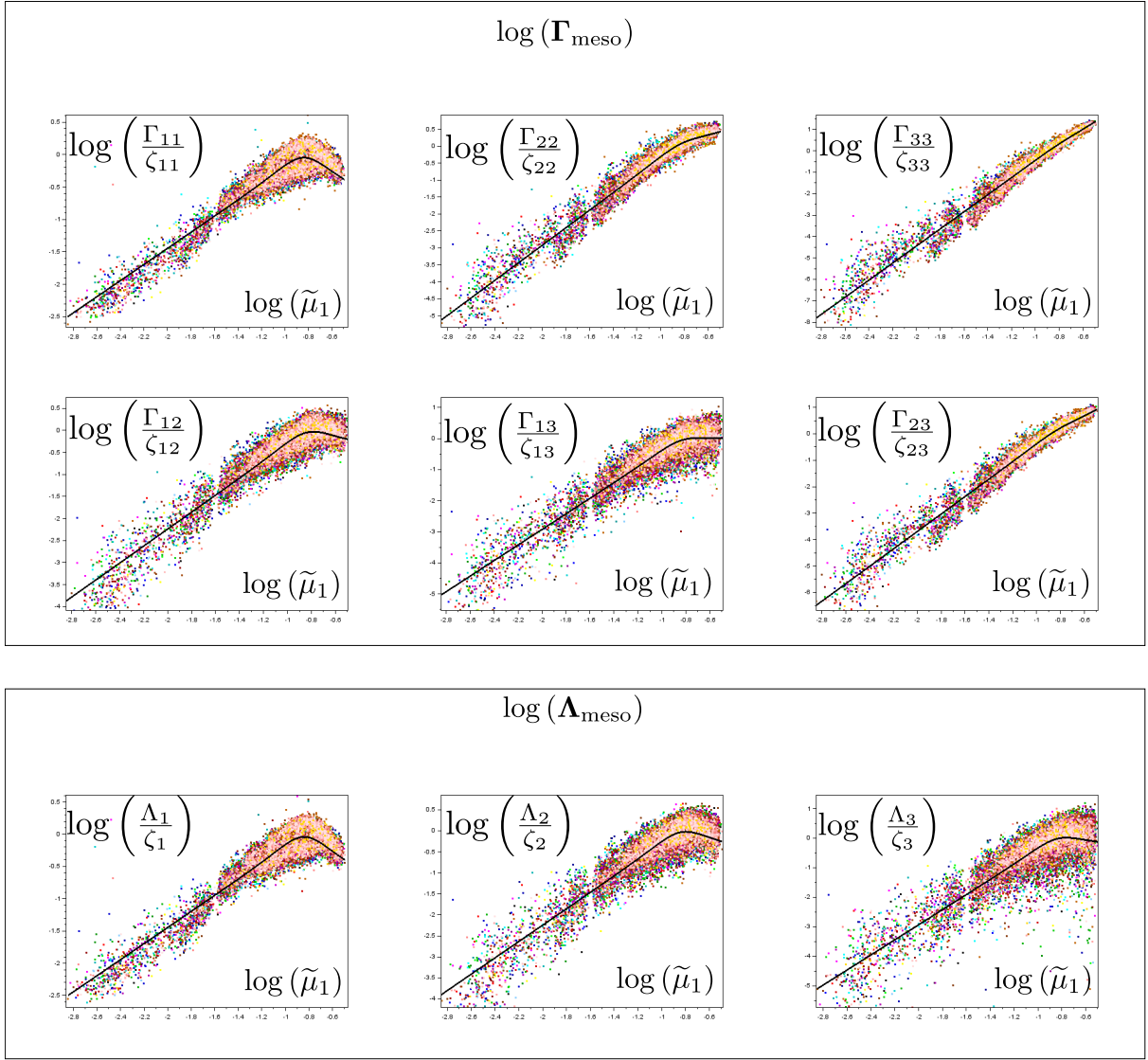


Figure 8: Rescaled data corresponding to  $\Gamma_{\text{meso}}$  and  $\Lambda_{\text{meso}}$  as a function of  $\tilde{\mu}_1$  computed from  $\mu_{\text{meso}}$  in logarithmic scales. Colors correspond to the tessellation index  $i_d$  ( $1 \leq i_d \leq N_{\text{tess}}$ ).

Table 4: Proportionality coefficients for  $\zeta_{(\cdot)}$ .

Component	Coefficient
11	2.189
22	1.431
33	1.000
12	1.019
13	0.543
23	1.137
1	61.980
2	28.428
3	14.450



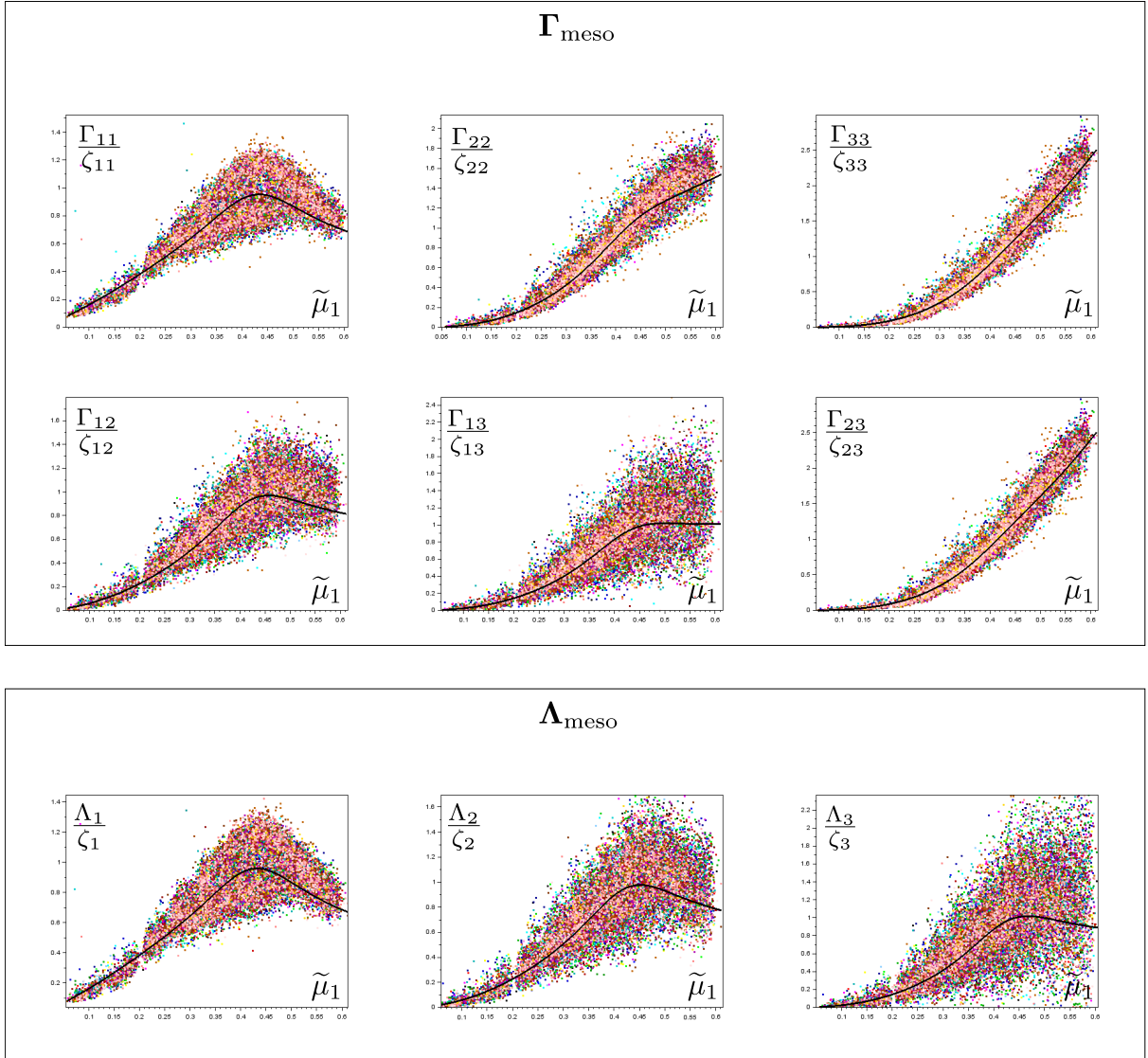


Figure 9: Raw data corresponding to  $\underline{\Gamma}_{\text{meso}}$  and  $\underline{\Lambda}_{\text{meso}}$  as a function of  $\tilde{\mu}_1$  computed from  $\underline{\mu}_{\text{meso}}$ . Colors correspond to the tessellation index  $i_d$  ( $1 \leq i_d \leq N_{\text{tess}}$ ).

It should be mentioned that the proposed linear model explains a large part of the variance observed in the data (i.e.,  $r^2 = 0.79$ ), however the linear relationship between the different components of  $\zeta_{(\cdot)}$  is imperfect (see. figure 10). Therefore, further work should focus on improving this particular aspect of the model.

A simple linear regression can be performed by minimizing with the least-squares method. However since there is a limited amount of data with non negligible statistical dispersion, it is preferable to identify the linear model within the framework of Bayesian inference as detailed in section 8.

In addition, since  $\eta_{\text{meso}}$  has a macroscopic counterpart  $\eta_{\text{macro}}$  defined in section 3.2, whose evolution has been obtained in (35), a macroscopic amplification factor can be defined on the basis of the interpolation function  $G_{(\cdot)}(\eta_{\text{macro}})$ .

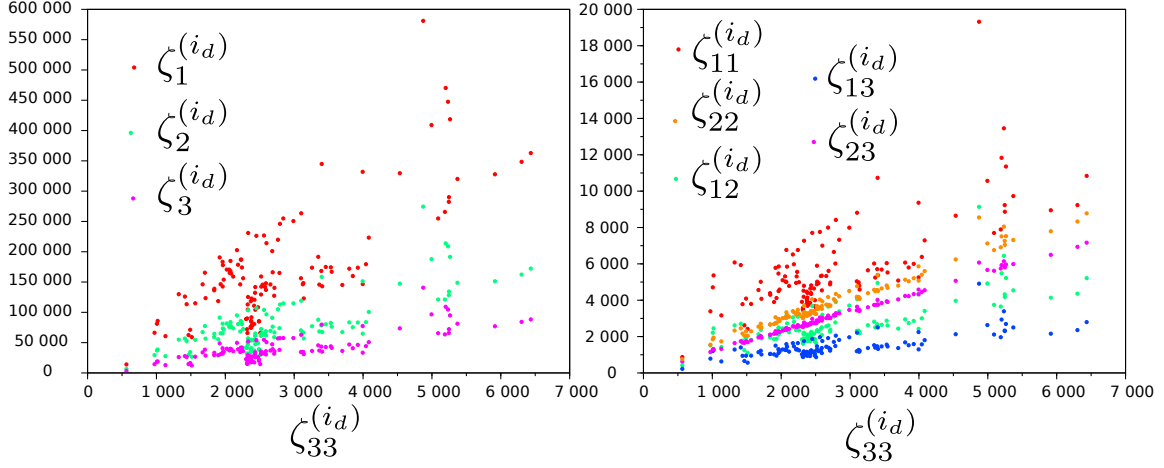


Figure 10: Correlation between the different components of  $\zeta_{(\cdot)}$

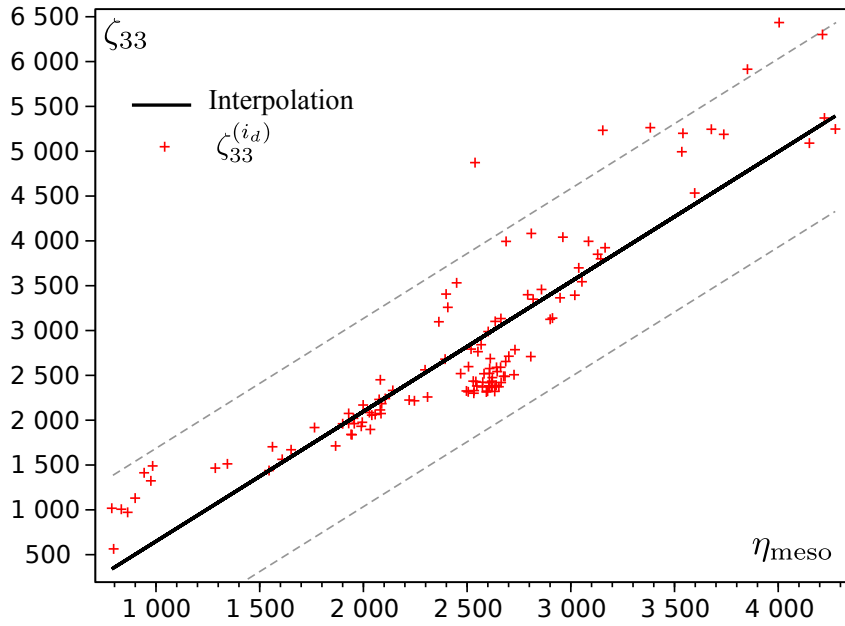


Figure 11: Amplification factor  $\zeta_{33}$  as a function of  $\eta_{meso}$

## 8. Epistemic uncertainty and Bayesian identification

### 8.1. Model for the tessellation amplification factor

In figures 9 and 11, it is clear that data are affected by a non negligible statistical dispersion around the interpolation models (61) and (63) respectively. Since the mesoscopic model introduced in [2] is fully deterministic, this statistical dispersion arises from the reduction of the amount of data introduced in section 3. Indeed, to a single overall state  $\mu_{meso}, \eta_{meso}$  correspond several different detailed mesoscopic states  $\alpha = (\mathbf{x}, \mathbf{y}, \mathbf{w})$ . The statistical dispersion evidenced in figures 9 and 11 shows that these different detailed mesoscopic states are not equivalent with respect to the mobility tensor  $\underline{\Gamma}_{meso}$ , the vector  $\underline{\Lambda}_{meso}$  and the amplification factor  $\zeta_{(\cdot)}$ . In other words, different RVEs sharing the same initial overall state  $\mu_{meso}, \eta_{meso}$  do not have the same evolution paths. Thus, the dispersion around the interpolation models (61) and (63) can be understood as the evidence of an *epistemic uncertainty* due to the loss of information in

the process of reducing the amount of data by introducing the overall state  $\mu_{\text{meso}}, \eta_{\text{meso}}$ . This epistemic uncertainty is not due to the rescaling process detailed in section 7.2. Indeed, the statistical dispersion can be observed for each tessellation individually as shown in figure 7. Thus, even for the same tessellation, the detailed arrangement of crystal orientations is lost in the overall state<sup>2</sup>, and the epistemic uncertainty arises.

As mentioned in section 7.4, parameters  $\widehat{A}, \widehat{B}$  involved in the interpolation model (63) are identified by using Bayesian techniques. It consists in considering a prior probability density for the model parameters  $\widehat{A}, \widehat{B}$ , and the likelihood, which is the conditional probability density of the data  $\zeta_{33}$  with respect to the model parameters  $\widehat{A}, \widehat{B}$ . Then, by using the Bayes theorem, the posterior probability density is obtained as the conditional probability of the model parameters with respect to the observed data. A normal model is considered, and the likelihood reads:

$$\zeta_{33}^{(id)} | \widehat{A}, \widehat{B}, \widehat{\sigma} \sim \mathcal{N}(\widehat{A} \eta_{\text{macro}} + \widehat{B}, \widehat{\sigma}^2) \quad (64)$$

where  $\mathcal{N}(\widehat{A} \eta_{\text{macro}} + \widehat{B}, \widehat{\sigma}^2)$  is the normal distribution of mean  $\widehat{A} \eta_{\text{macro}} + \widehat{B}$ , and  $\widehat{\sigma}$  is the standard deviation, which characterizes the dispersion around the mean. In addition, Bayesian techniques rely on prior distributions for the unknown parameters. An informative normal distribution is chosen for  $\widehat{A}$  as the slope can be clearly identified in figure 11, and a non-informative uniform distribution is chosen for  $\widehat{B}$ . Classically conjugate standard deviation prior follows an inverted chi-squared probability density. Hence:

$$\begin{cases} \widehat{A} \sim \mathcal{N}(\mu_{\widehat{A}}, s_{\widehat{A}}) \\ \widehat{B} \sim \mathcal{U}(\widehat{B}_{\min}, \widehat{B}_{\max}) \\ \widehat{\sigma} \sim \text{Inv-}\chi^2(\widehat{s}^2, \widehat{\nu}) \end{cases} \quad (65)$$

where parameters  $\mu_{\widehat{A}}, s_{\widehat{A}}, \widehat{B}_{\min}, \widehat{B}_{\max}, \widehat{s}, \widehat{\nu}$  are chosen so that the priors are consistent with figure 11:

$$\begin{cases} \mu_{\widehat{A}} = 1.5 \\ s_{\widehat{A}} = 0.3 \\ \widehat{B}_{\min} = -1300 \\ \widehat{B}_{\max} = 300 \\ \widehat{s} = 400 \\ \widehat{\nu} = 20 \end{cases} \quad (66)$$

Bayesian inference consists in computing the posterior probability density by using the Bayes theorem, hence:

$$p(\widehat{A}, \widehat{B}, \widehat{\sigma} | \zeta_{33}) \propto p(\zeta_{33} | \widehat{A}, \widehat{B}, \widehat{\sigma}) p(\widehat{A}) p(\widehat{B}) p(\widehat{\sigma}) \quad (67)$$

Statistics of posterior probability density function (67) are explored by Markov-Chain Monte Carlo (MCMC) sampling techniques. In practice, a No U-Turn Sampler (NUTS) developed by [49] is used within the framework of the PYMC3 package developed by [50] in PYTHON ([51]). The posterior densities are presented in figure 12 as well as a scatter-plot of pairwise joint densities showing a correlation between  $\widehat{A}$  and  $\widehat{B}$ . In addition, maximum a posteriori estimates are computed and listed in table 5, and have been used in figure 11.

---

<sup>2</sup>even though a part of the information is not lost as mean and standard deviation of the BMD are managed at the macroscopic scale

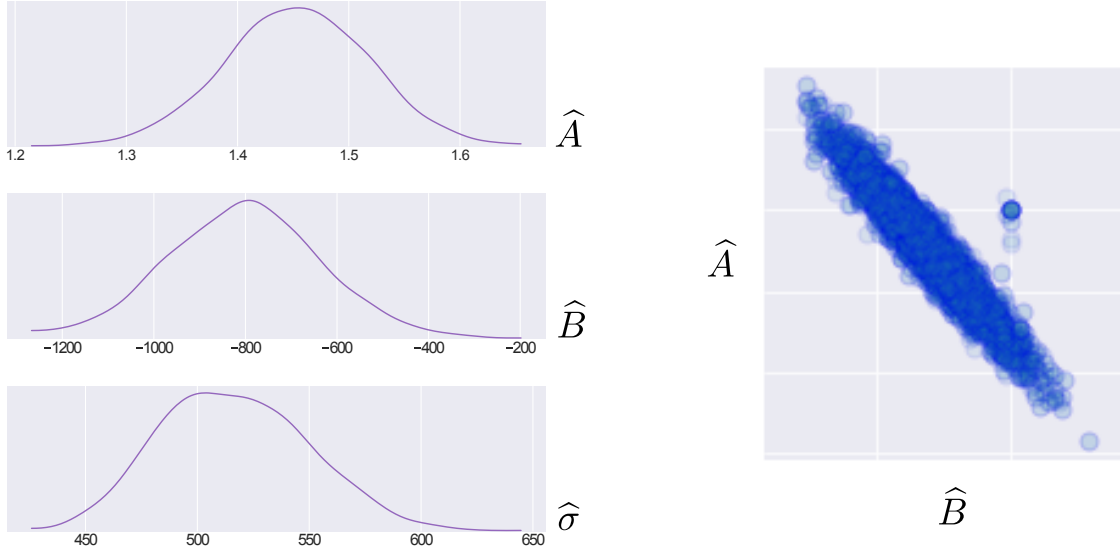


Figure 12: Posterior densities and pairplot

Table 5: Maximum a posteriori estimates.

Component	$\widehat{A}$	$\widehat{B}$	$\widehat{\sigma}$
33	1.451	-802	512

Thus, one can define a macroscopic deviation with respect to the interpolation model (63) denoted by  $\widehat{\epsilon}_{\text{macro}}$ , whose probability density function reads:

$$\widehat{\epsilon}_{\text{macro}} \sim \mathcal{N}(0, \widehat{\sigma}^2) \quad (68)$$

## 8.2. Uncertainty associated to the macroscopic mobility

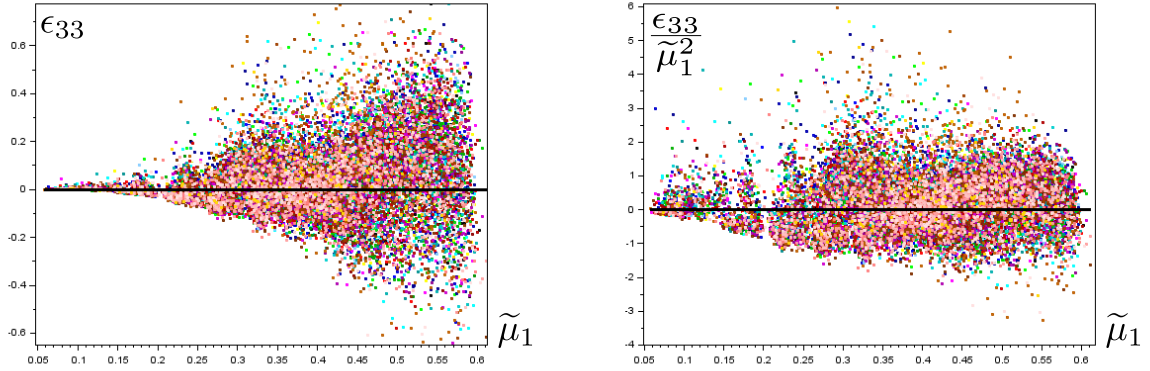
For each entry  $k$  in the rescaled database (60), the deviation with respect to the interpolation model (61) is denoted by  $\epsilon_{(\cdot)}^{(k)}$  and reads:

$$\forall i_d \in \{1, \dots, N_{\text{tess}}\}, \forall k \in \mathcal{K}(i_d), \epsilon_{(\cdot)}^{(k)} = \frac{\Gamma_{(\cdot)}^{(k)}}{\zeta_{(\cdot)}^{(i_d)}} - \exp \left[ F_{(\cdot)}^{\text{sc}} \left( \log \left( \widetilde{\mu}_1^{(k)} \right) \right) \right] \quad (69)$$

Of course, replacing  $\Gamma_{(\cdot)}^{(k)}$  by  $\Lambda_{(\cdot)}^{(k)}$  in (69), the same definition applies for the deviation with respect to the interpolation model of  $\Lambda_{\text{meso}}$ . For instance  $\epsilon_{33}^{(k)}$  is presented as a function of  $\widetilde{\mu}_1^{(k)}$  for  $1 \leq k \leq N_{\text{data}}$  in figure 13a. It is clear that the deviation depends on  $\widetilde{\mu}_1$ , and  $\epsilon_{33}^{(k)} / \left( \widetilde{\mu}_1^{(k)} \right)^2$  is presented in figure 13b. The corresponding histogram is presented in figure 15 showing that a Gaussian model is acceptable.

Since the different components of the deviation (69) are correlated, a multivariate Gaussian distribution is chosen. Thus, one can define the macroscopic deviation as a random variable vector denoted by  $\epsilon_{\text{macro}} / \widetilde{\mu}_1^2$  whose probability density function is a multivariate Gaussian distribution of zero mean and covariance matrix  $\underline{\mathbf{C}}$  of size  $9 \times 9$ . One can estimate the probability density function of  $\underline{\mathbf{C}}$  within the framework of Bayesian inference. Since a Gaussian model has been chosen, the likelihood is the following conditional probability density function:

$$\frac{\epsilon_{\text{macro}}}{\widetilde{\mu}_1^2} | \underline{\mathbf{C}} \sim \mathcal{N}_9(0, \underline{\mathbf{C}}) \quad (70)$$



(a)  $\epsilon_{33}$  vs  $\tilde{\mu}_1$ .

(b)  $\epsilon_{33}/\tilde{\mu}_1^2$  vs  $\tilde{\mu}_1$ .

Figure 13: Statistical dispersion around the interpolation function.

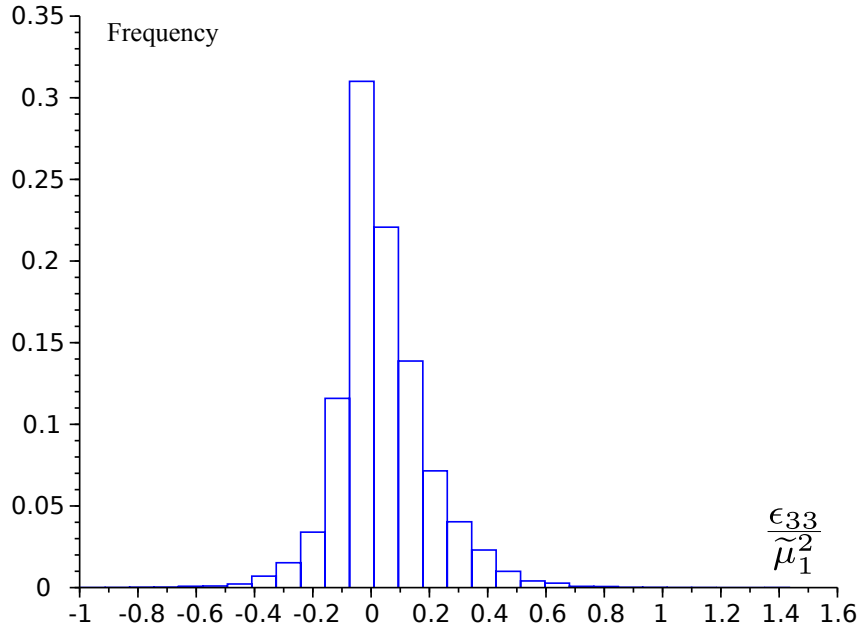


Figure 14: Histogram of statistical dispersion around the interpolation function for the 33 component of  $\underline{\Gamma}_{meso}$ .

Figure 15: Statistical dispersion around the interpolation function.

The inverse-Wishart distribution is the conjugate prior distribution for the covariance matrix of a multivariate normal distribution. However, such a distribution is not implemented in usual Bayesian computational packages. Thus, the LKJ-Cholesky covariance prior is used, which is well-suited to computational Bayesian methods:

$$\underline{\mathbf{L}} \sim \text{LKJ-C}(9, \nu, \sigma) \quad (71)$$

where  $\underline{\mathbf{L}}$  is a lower triangular matrix arising in the Cholesky decomposition of  $\underline{\mathbf{C}}$ , namely:

$$\underline{\mathbf{C}} = \underline{\mathbf{L}} \cdot \underline{\mathbf{L}}^T \quad (72)$$

and where  $\text{LKJ-C}(9, \nu, \sigma)$  the the LKJ-Cholesky distribution of size  $9 \times 9$ ,  $\nu$  is a parameter, which controls the amount of correlation between components of  $\epsilon_{macro}/\mu_1^2$ , and  $\sigma$  controls the

distribution of standard deviation. The following parameters are chosen for the prior distribution:

$$\begin{cases} \sigma \sim \mathcal{E}(\lambda) \\ \lambda = 1 \\ \nu = 2 \end{cases} \quad (73)$$

where  $\mathcal{E}(\lambda)$  is the exponential density function of parameter  $\lambda$ . The posterior density function reads:

$$p\left(\underline{\mathbf{L}} \mid \frac{\boldsymbol{\epsilon}_{\text{macro}}}{\widetilde{\mu}_1^2}\right) \propto p\left(\frac{\boldsymbol{\epsilon}_{\text{macro}}}{\widetilde{\mu}_1^2} \mid \underline{\mathbf{C}}\right) p(\underline{\mathbf{L}}) \quad (74)$$

Statistics of posterior probability density function (74) are explored by Markov-Chain Monte Carlo (MCMC) sampling techniques. In practice, a No U-Turn Sampler (NUTS) developed by [49] is used within the framework of the PYMC3 package developed by [50] in Python ([51]).

### 8.3. Time dependance and numerical implementation

The fundamental nature of the *epistemic uncertainty* can be understood as follows. Consider random variables denoted by  $\Omega_{(\cdot)}$ , each of which corresponds to a component ( $\cdot$ ) of  $\underline{\Gamma}_{\text{meso}}$  and  $\underline{\Lambda}_{\text{meso}}$ . The random variable vector  $\boldsymbol{\Omega} = (\Omega_{11}, \dots, \Omega_{23}, \Omega_1 \dots, \Omega_3)$  of size 9 captures the uncertainty due to the random “selection” of a virtual mesoscopic state given a macroscopic state. Ideally, a single scalar random variable  $\Omega^*$  could account for this random “selection”. However, the link between the different components of  $\underline{\Gamma}_{\text{meso}}$  and  $\underline{\Lambda}_{\text{meso}}$ , which would enable from a single draw of  $\Omega^*$  to compute the deviations associated to the different components, is unknown at the macroscopic scale and fully depends on the mesoscopic scale. That is why at the macroscopic scale the uncertainty is modeled by a random variable vector  $\boldsymbol{\Omega}$ . It should be noted that the covariance matrix  $\underline{\mathbf{C}}$  represents the macroscopic link between the different components. The deviation with respect to the interpolation model reads:

$$\begin{cases} \boldsymbol{\epsilon}_{\text{macro}} = \widetilde{\mu}_1^2 \underline{\mathbf{L}} \cdot \boldsymbol{\Omega} \\ \widehat{\boldsymbol{\epsilon}}_{\text{macro}} = \widehat{\sigma} \widehat{\boldsymbol{\Omega}} \end{cases} \quad (75)$$

where  $\underline{\mathbf{L}}$  is defined in (72). In addition,  $\Omega_{(\cdot)}$  are independent and identically distributed standard normal distributions as well as  $\widehat{\boldsymbol{\Omega}}$ :

$$\begin{cases} \Omega_{(\cdot)} \sim \mathcal{N}(0, 1) \\ \widehat{\boldsymbol{\Omega}} \sim \mathcal{N}(0, 1) \end{cases} \quad (76)$$

The macroscopic model is therefore probabilistic. However, since the mesoscopic model is fully deterministic, there is no fundamental indeterminacy, only an epistemic uncertainty within the upscaling process due to the loss of information.

At this stage the macroscopic counterparts  $\underline{\Gamma}_{\text{macro}}$  and  $\underline{\Lambda}_{\text{macro}}$  of  $\underline{\Gamma}_{\text{meso}}$  and  $\underline{\Lambda}_{\text{meso}}$  can be stated:

$$\begin{cases} \underline{\Gamma}_{\text{macro}} = (\underline{\mathbf{G}}(\eta_{\text{macro}}) + \widehat{\underline{\mathbf{v}}}_{\text{macro}}) (\exp[\underline{\mathbf{F}}^{\text{sc}}(\boldsymbol{\mu}_{\text{macro}})] + \underline{\mathbf{v}}_{\text{macro}}) \\ \underline{\Lambda}_{\text{macro}} = (\underline{\mathbf{G}}(\eta_{\text{macro}}) + \widehat{\underline{\mathbf{v}}}_{\text{macro}}) (\exp[\underline{\mathbf{F}}^{\text{sc}}(\boldsymbol{\mu}_{\text{macro}})] + \underline{\mathbf{v}}_{\text{macro}}) \end{cases} \quad (77)$$

where  $\underline{\mathbf{F}}^{\text{sc}}$  and  $\mathbf{F}^{\text{sc}}$  are the tensor and vector interpolation functions defined from (61),  $\underline{\mathbf{G}}$  and  $\mathbf{G}$  are the tensor and vector interpolation functions defined from (63). In addition  $\underline{\mathbf{v}}_{\text{macro}}$  and  $\mathbf{v}_{\text{macro}}$  are the tensors and vectors derived from  $\boldsymbol{\epsilon}_{\text{macro}}$  defined in (75). Second order tensors are

symmetric of size  $3 \times 3$  composed of the 11, 22, 33, 12, 13 and 23 components, and vectors are composed of the 1, 2, 3 components. Similarly  $\widehat{\mathbf{v}}_{\text{macro}}$  and  $\widehat{\mathbf{v}}_{\text{macro}}$  are obtained from  $\widehat{\epsilon}_{\text{macro}}$  defined in (68) and the proportionality coefficients listed in table 4.

With the macroscopic definition (77), the evolution law (28) for  $\boldsymbol{\mu}_{\text{macro}}$  and the evolution law (35) for  $\eta_{\text{macro}}$  can be computed numerically. The main difficulty lies in the fact that the random vector  $\boldsymbol{\Omega}$  correspond to the random “selection” of a virtual mesoscopic state for given a known macroscopic state. However, the mesoscopic state slowly evolves during grain growth. Thus, this “selection” of a virtual mesoscopic state should be updated at each time step. If new draws of the random vector  $\boldsymbol{\Omega}$  were computed at each time step, a completely new virtual mesoscopic state would be “selected” leading to Brownian motion. Of course, this does not correspond to the epistemic uncertainty that has been introduced. Ideally, at each time step new draws of  $\boldsymbol{\Omega}$ , which strongly depend on the draws at the previous time step, would be computed. However, this would need to identify an autocovariance matrix, which is the function of the covariance with respect to pairs of different time steps. In this work, the identification of the macroscopic model has been done according to a *static* approach, as mentioned in section 5. Thus, the identification of such an autocovariance matrix is not possible. To overcome this difficulty, we assume that the virtual mesoscopic state, which has been “selected” by the initial draw of the random vector  $\boldsymbol{\Omega}$ , evolves sufficiently slowly so that there is no need to “re-select” a new virtual mesoscopic state for further time steps. That is to say that  $\boldsymbol{\Omega}$  does not depend on time, and draws of  $\boldsymbol{\Omega}$  are computed only at  $t = 0$  and remain constant during the evolution. However, since  $\epsilon_{\text{macro}}$  in (75) explicitly depends on  $\widetilde{\mu}_1(t)$  (computed from  $\boldsymbol{\mu}_{\text{macro}}$ ), the overall epistemic uncertainty is nevertheless adjusted during the evolution, but in a simplified way.

This leads to consider the evolution laws (28) and (35) as simple stochastic processes. The evolutions of means  $\langle \boldsymbol{\mu}_{\text{macro}} \rangle(t)$ ,  $\langle \eta_{\text{macro}} \rangle(t)$  and point-wise standard deviations  $\sigma_{\boldsymbol{\mu}_{\text{macro}}}(t)$ ,  $\sigma_{\eta_{\text{macro}}}(t)$  can be computed easily by simulating a large number of evolutions with various draws of  $\boldsymbol{\Omega}$ ,  $\widehat{\epsilon}_{\text{macro}}$ . One could significantly reduce the number of evolutions to accurately estimate  $\langle \boldsymbol{\mu}_{\text{macro}} \rangle(t)$ ,  $\langle \eta_{\text{macro}} \rangle(t)$  and  $\sigma_{\boldsymbol{\mu}_{\text{macro}}}(t)$ ,  $\sigma_{\eta_{\text{macro}}}(t)$  by using spectral approaches [52], but this has not been done in this study as computation time is fairly reduced.

## 9. Results

In this section, the macroscopic model is implemented numerically and applied to various conditions. More precisely the evolution laws (28) and (35) are discretized with a simple explicit scheme. It should be mentioned that the temperature  $T$  in (28) and (35) is an external quantity, which of course can depend on time if temperature cycles are considered for instance. To take into account time dependent temperature it is sufficient to update the temperature at each time step in the explicit scheme. However, in the following constant temperature is considered.

Comparisons with the evolution of the overall mesoscopic state are provided to evaluate the model quality. Several tessellations included in the data base  $i_d \in \mathcal{I}$  where  $\mathcal{I} = \{1, 80, 100\}$  have been considered as initial states for mesoscopic evolutions with crystal orientations assigned with a beta probability density  $\mathcal{B}(\alpha, \beta)$  where  $(\alpha, \beta) = (1, 1)$  and  $(\alpha, \beta) = (5, 9)$ . These tessellations have been selected so that very different morphological textures can be tested. In addition, four tessellations that do not belong to the database have also been considered with very similar initial overall mesoscopic state (these four tessellations are referred in the following as  $i_d = 0$ ). All mesoscopic evolutions have been computed according to the mesoscopic model detailed in [2], and parameters are listed in table 6. Some examples of OTs at different

times during grain growth are presented in figures 16, 17 and 18.

Table 6: Simulation parameters used for the mesoscopic evolutions.

Temperature	(°C)	$T$	800
Mobility	( $\text{m}^4 \cdot \text{J}^{-1} \cdot \text{s}^{-1}$ )	$m(T)$	$0.146 \times 10^{-13}$
Physical size	(mm)	$L_0$	1
Duration	(hours)	$d$	10

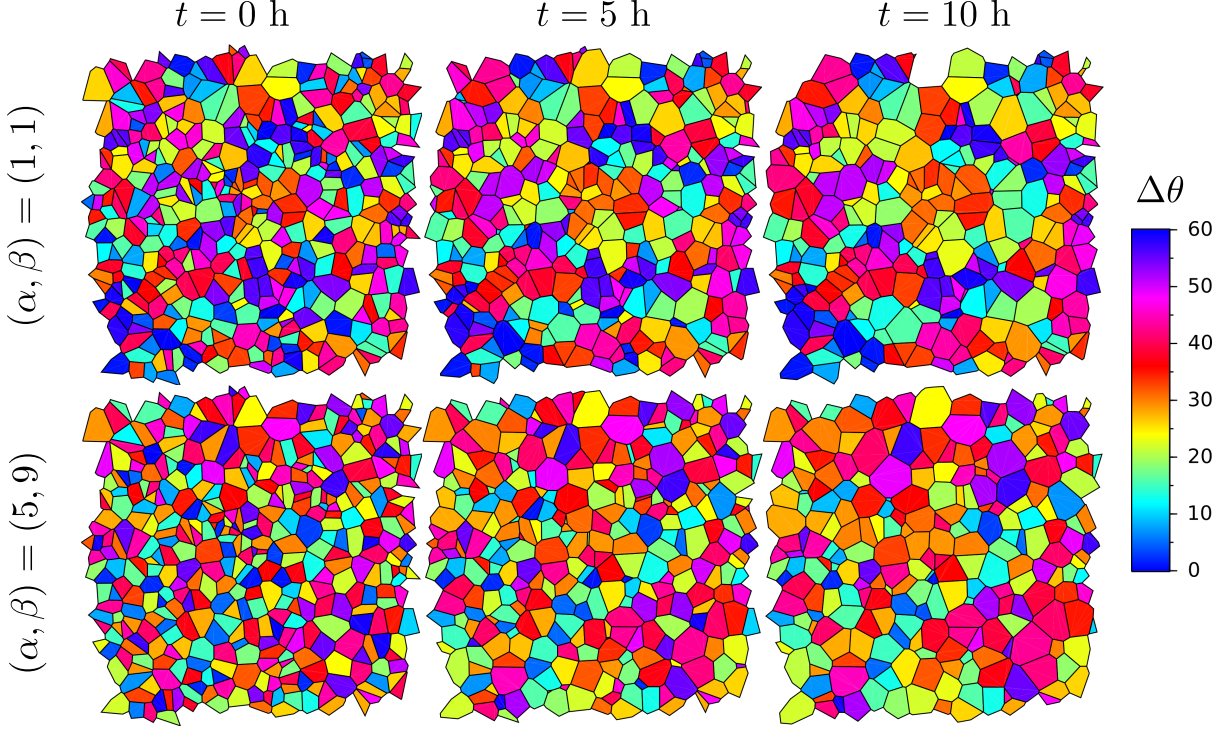


Figure 16: OTs at different times for two different BMDs obtained with a beta density  $\mathcal{B}(\alpha, \beta)$  for  $(\alpha, \beta) = (1, 1)$  and  $(\alpha, \beta) = (5, 9)$ ,  $i_d = 1$ .

### 9.1. First validation and numerical implementation

Since the macroscopic model is probabilistic, typical results are the means  $\langle \mu_{\text{macro}} \rangle(t)$ ,  $\langle \eta_{\text{macro}} \rangle(t)$  and point-wise standard deviations  $\sigma_{\mu_{\text{macro}}}(t)$ ,  $\sigma_{\eta_{\text{macro}}}(t)$ . However, a simplified condition is considered in this section, and consists in neglecting the epistemic uncertainty, that is to say neglecting  $\underline{\mathbf{v}}_{\text{macro}}$ ,  $\widehat{\mathbf{v}}_{\text{macro}}$ ,  $\mathbf{v}_{\text{macro}}$ , and  $\widehat{\mathbf{v}}_{\text{macro}}$  in (77). Thus, a single particular macroscopic evolution is compared to the evolution of the overall mesoscopic state  $\mu_{\text{meso}}, \eta_{\text{meso}}$  in order to determine how both models behave with respect to smoothness.

Comparisons for  $i_d = 1$  are presented in figure 19. These results show that the general behavior of the macroscopic model is satisfying, but as expected it is smoother than the overall mesoscopic evolution. At the beginning of the evolution the macroscopic model is rather accurate even though it may significantly diverge from the overall mesoscopic behavior after the first few hours. A perfect match is of course not expected as the macroscopic evolution is a particular draw of a stochastic process. Nevertheless, the proposed comparisons show that the macroscopic model is unable to capture sudden slope changes for long duration simulations.



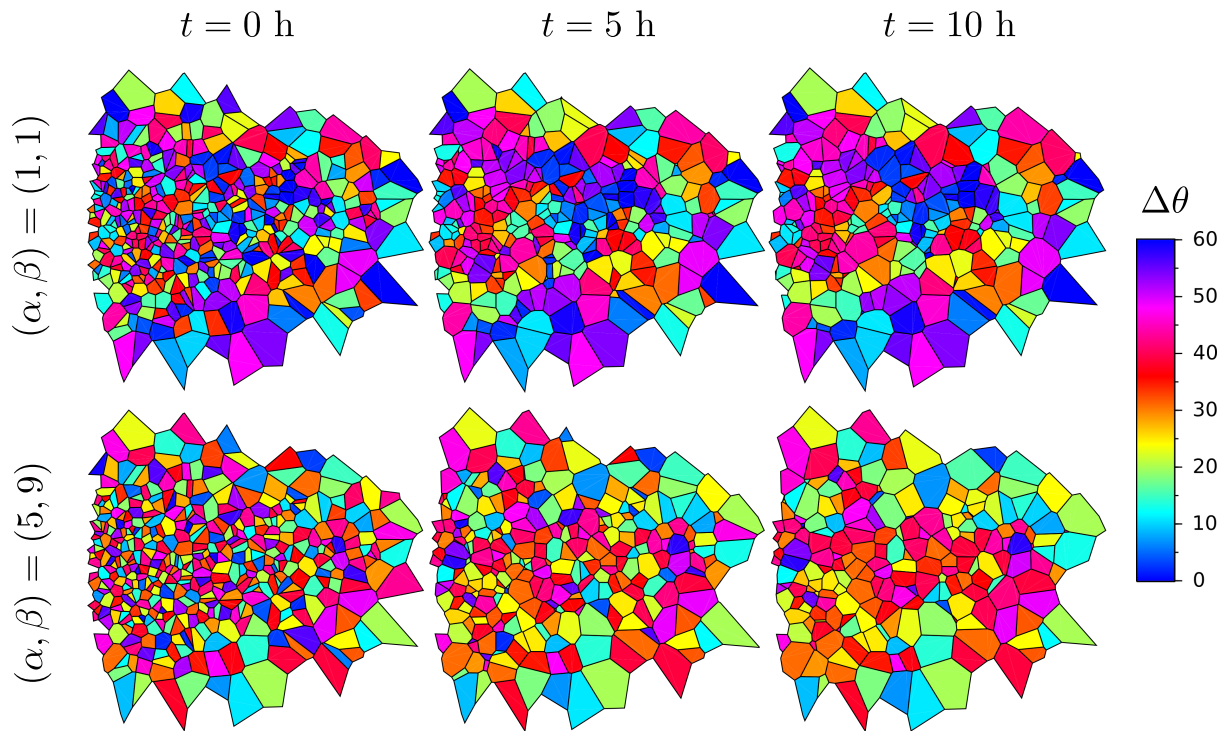


Figure 17: OTs at different times for two different BMDs obtained with a beta density  $\mathcal{B}(\alpha, \beta)$  for  $(\alpha, \beta) = (1, 1)$  and  $(\alpha, \beta) = (5, 9)$ ,  $i_d = 80$ .

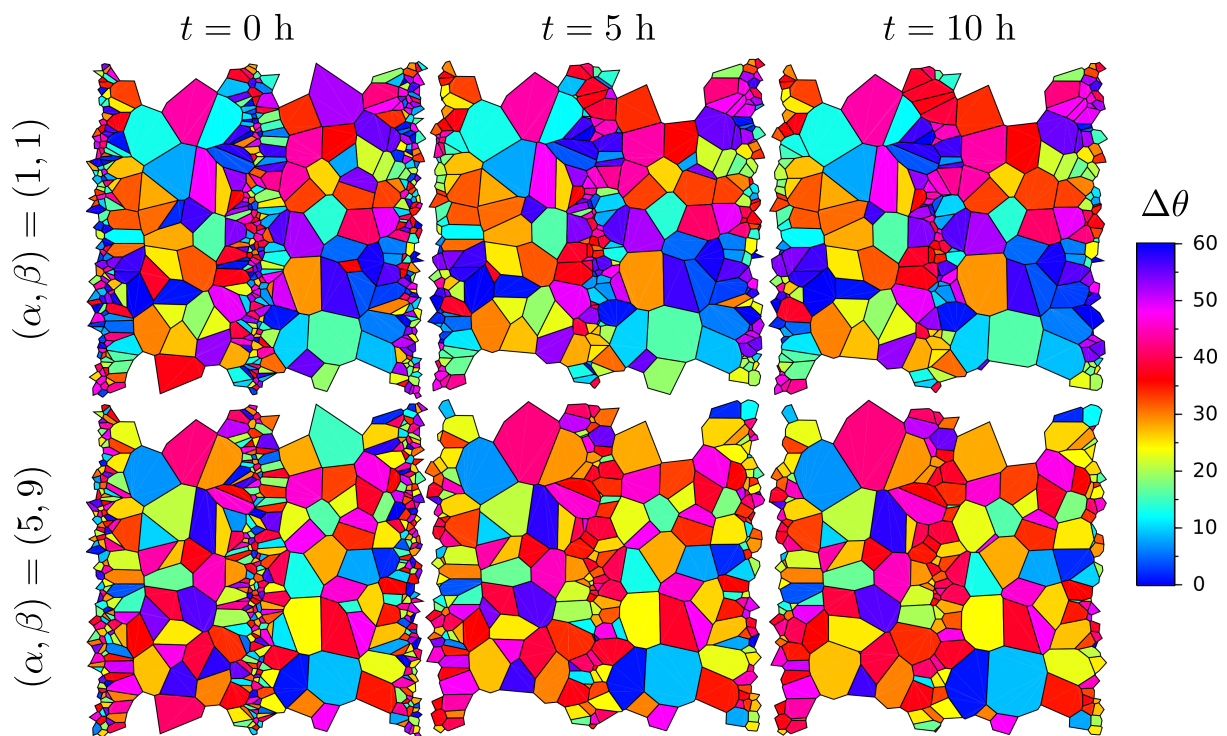


Figure 18: OTs at different times for two different BMDs obtained with a beta density  $\mathcal{B}(\alpha, \beta)$  for  $(\alpha, \beta) = (1, 1)$  and  $(\alpha, \beta) = (5, 9)$ ,  $i_d = 100$ .

These slope changes correspond to local events that have significant overall effect. However,

$\underline{\Gamma}_{\text{macro}}$  depends on  $\mu_{\text{macro}}, \eta_{\text{macro}}$  in a smooth manner, which explains the regularity of the macroscopic evolutions. This observation shows that the proposed *static* approach, which leads to set the epistemic uncertainty at a constant value during the evolution, contributes to the smoothness of the obtained solution. To take into account sudden slope changes at the macroscopic scale one could propose a *dynamic* approach with a probability of local “events” increasing with time, and leading to sudden changes of  $\underline{\Gamma}_{\text{macro}}$ . However, such an approach seems difficult to develop as the identification of the probability density of local “events” would necessitate to analyze in details a large number of mesoscopic evolutions, although the database used in this study is only composed of various and disconnected mesoscopic states.

## 9.2. Probabilistic results

In this section, the epistemic uncertainty is not neglected in (77), thus the probabilistic nature of the macroscopic model is evidenced. Draws of the various model parameters involved in the Bayesian identification are obtained as detailed in section 8 and used to generate  $N_{\text{stat}} = 500$  macroscopic evolutions whose means, medians and point-wise standard deviations are computed. The comparisons between the overall mesoscopic state  $\mu_{\text{meso}}, \eta_{\text{meso}}$  and the mean macroscopic state  $\langle \mu_{\text{macro}} \rangle (t), \langle \eta_{\text{macro}} \rangle (t)$  with point-wise standard deviations  $\sigma_{\mu_{\text{macro}}} (t), \sigma_{\eta_{\text{macro}}} (t)$  are presented for the different tested conditions in figures 20, 21 and 22 for  $i_d \in \{1, 80, 100\}$  and in figure 23 for the 4 tessellations, which do not belong to the database (i.e.,  $i_d = 0$ ). The zone defined by more or less one standard deviation is centered on the median and not the mean in case the distribution is not symmetrical (e.g.,  $\tilde{\mu}_2$ ). Results show that the macroscopic model is satisfying, as mesoscopic evolutions lie for the most part in the zone defined by more or less one standard deviation around the median.

It should be noted that the mesoscopic “events”, which are not easily captured by the macroscopic model are more likely for tessellations whose morphological textures are very pronounced (i.e.,  $i_d = 80$  or  $i_d = 100$ ). This is a general trend that has been checked for many other tessellations. Nevertheless, the macroscopic model proposed in this study seems sufficiently accurate to carry out macroscopic simulations of annealing processes or large fabrication or forming processes with heterogeneous and unsteady temperature.

In figure 23 all the four mesoscopic evolutions are stacked in one half of the zone defined by the point-wise standard deviation for  $\mu_1, \mu_2$ . This observation could raise the idea that there is a bias, that is to say a systematic error in the estimation of the mean evolution. However, this is likely due to the fact that all the four evolutions have been obtained from similar tessellations to facilitate the search of almost identical initial overall mesoscopic states  $\mu_{\text{meso}}(t = 0), \eta_{\text{meso}}(t = 0)$ . Since the macroscopic model has been identified by using very different tessellations, the epistemic uncertainty reflects this diversity, and the macroscopic results spread on a larger zone than if similar tessellations were used for the identification. Thus, for some applications, if equiaxed grains are the most likely mesoscopic structure, the macroscopic model could be identified with a specific database, which only involves such tessellations, and therefore the epistemic uncertainty could be reduced.

## 9.3. Average grain size

Results obtained in section 9.2 provide interesting information such as the GB length density per unit area  $\mu_0/L_0$  (m<sup>-1</sup>), the average misorientation  $\tilde{\mu}_1$  (rad), the square of the relative standard deviation  $\tilde{\mu}_2$ , and the standard deviation that can be obtained from  $\tilde{\mu}_1, \tilde{\mu}_2$ . In addition to these statistical descriptors of the polycrystalline structure, a morphological quantity is also

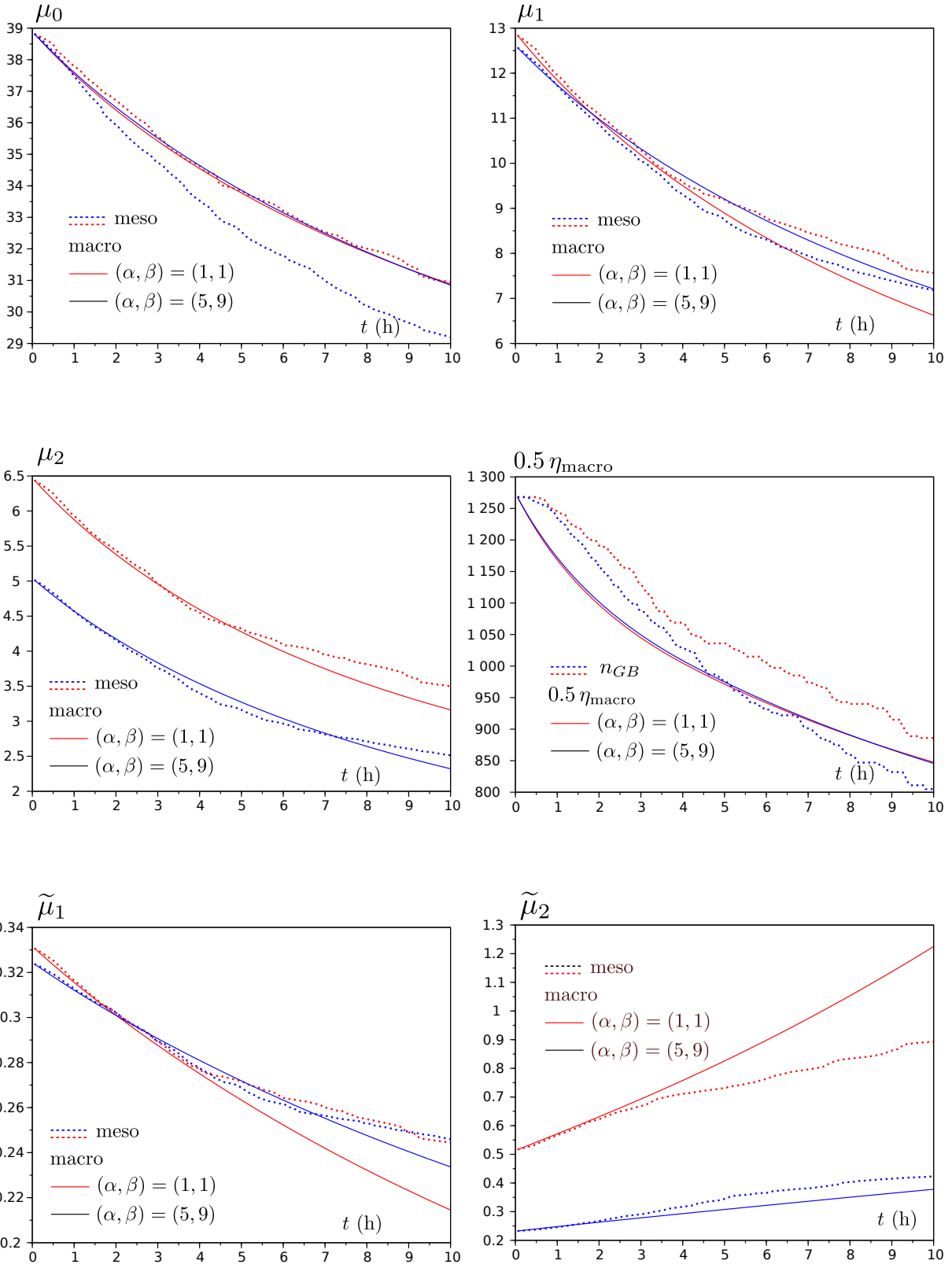


Figure 19: Comparison between the overall mesoscopic state and one particular macroscopic state obtained by neglecting the epistemic uncertainty,  $i_d = 1$ .

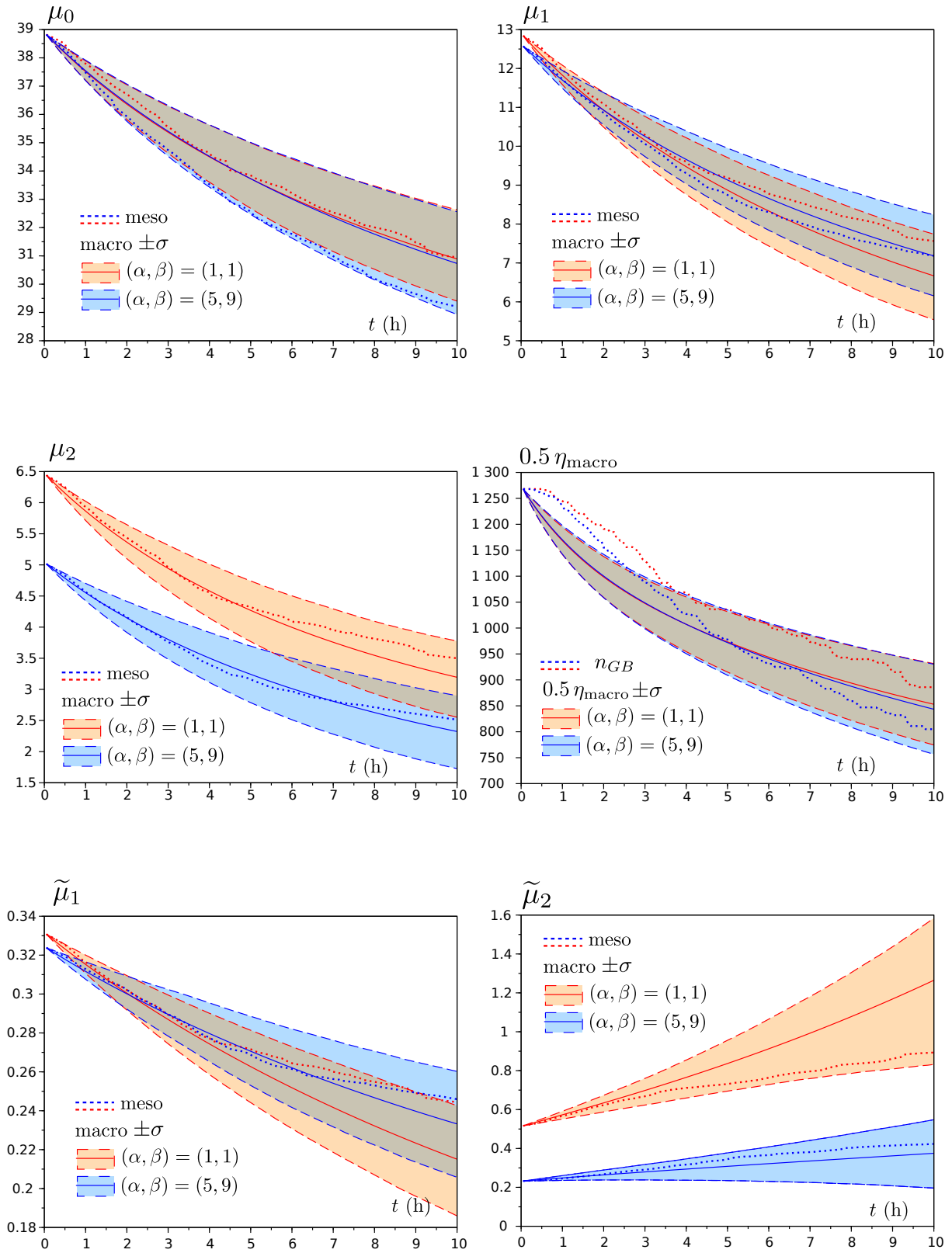


Figure 20: Comparison between overall mesoscopic state and macroscopic state with point-wise standard deviation  $\pm\sigma$ ,  $i_d = 1$ .

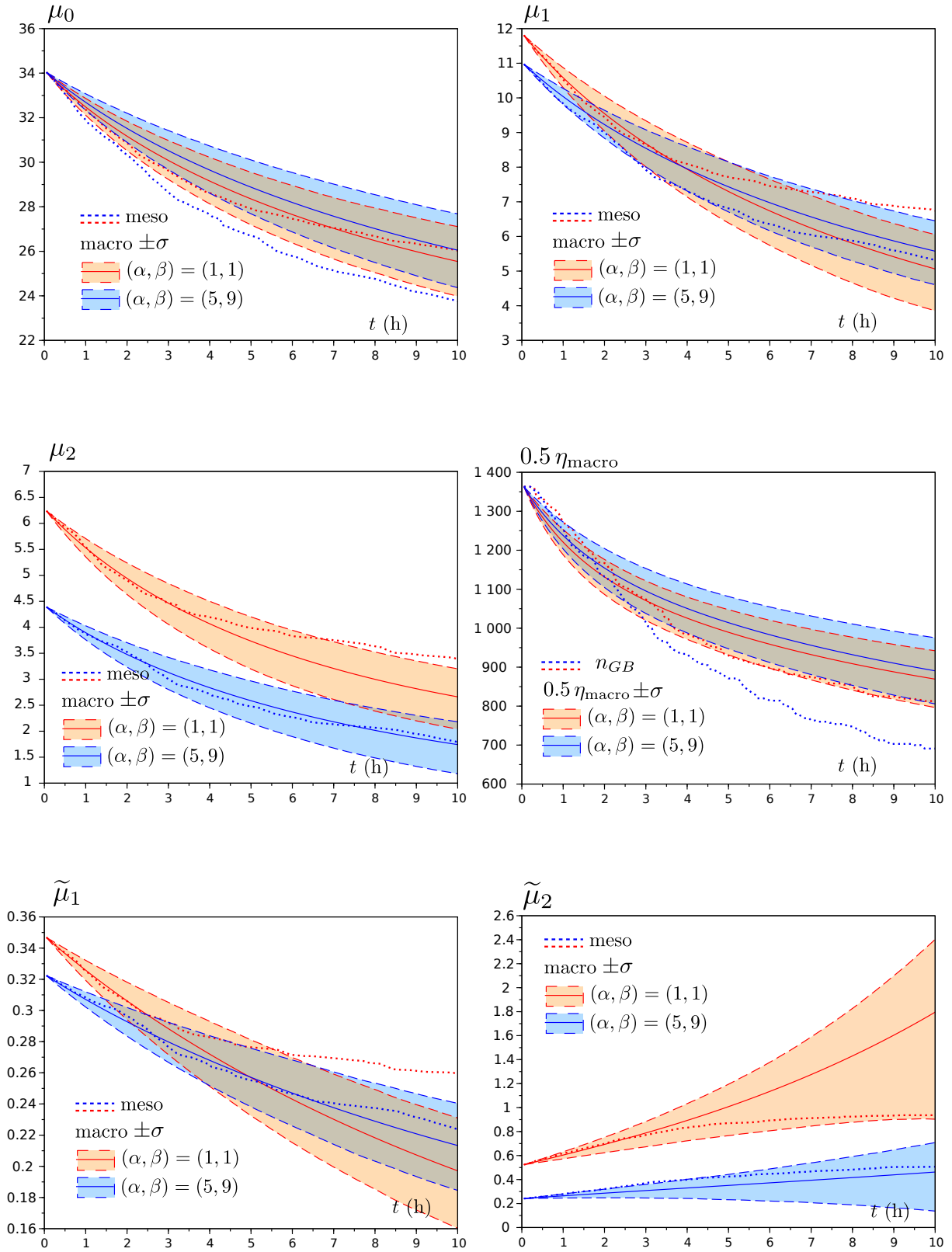


Figure 21: Comparison between overall mesoscopic state and macroscopic state with point-wise standard deviation  $\pm\sigma$ ,  $i_d = 80$ .

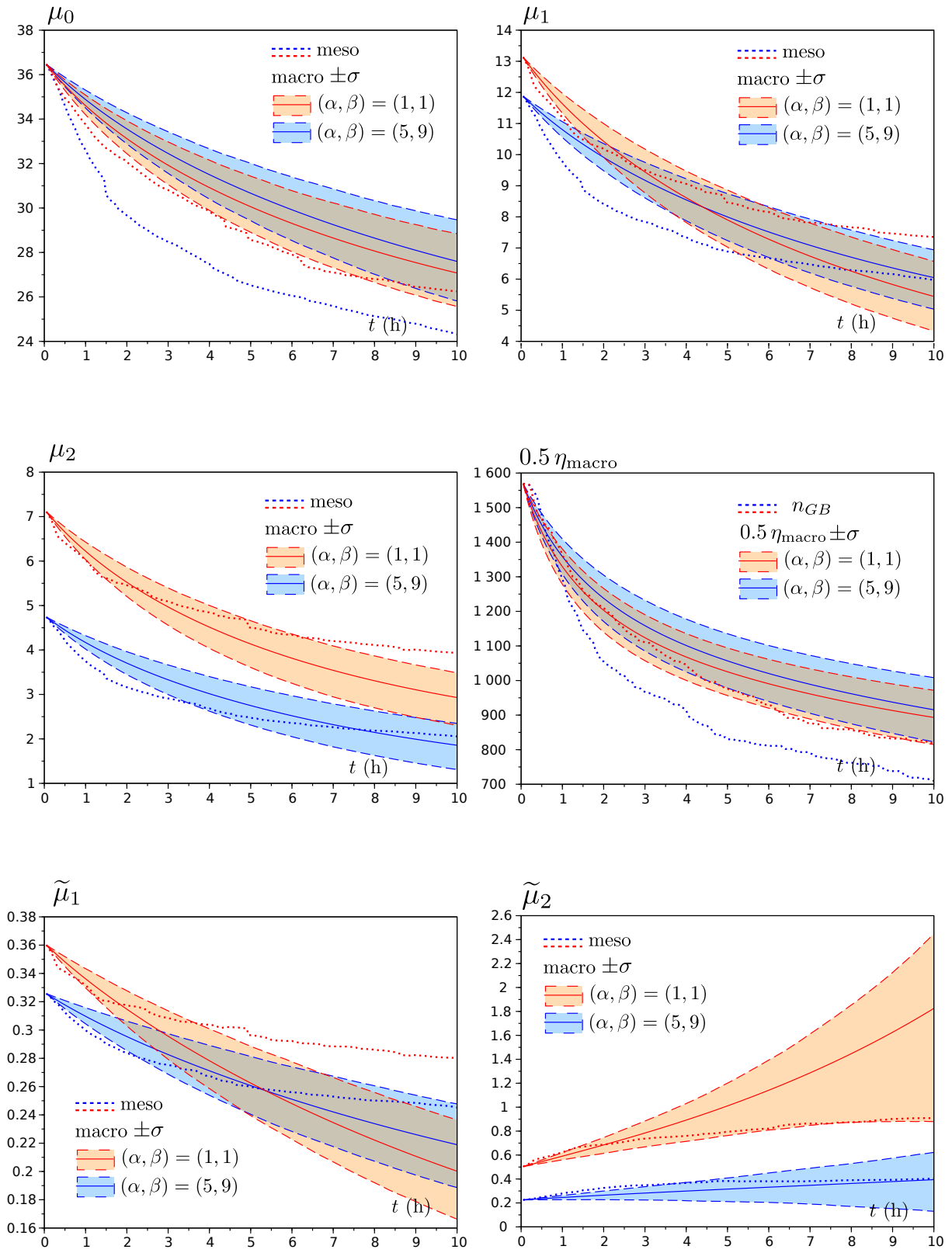


Figure 22: Comparison between overall mesoscopic state and macroscopic state with point-wise standard deviation  $\pm\sigma$ ,  $i_d = 100$ .

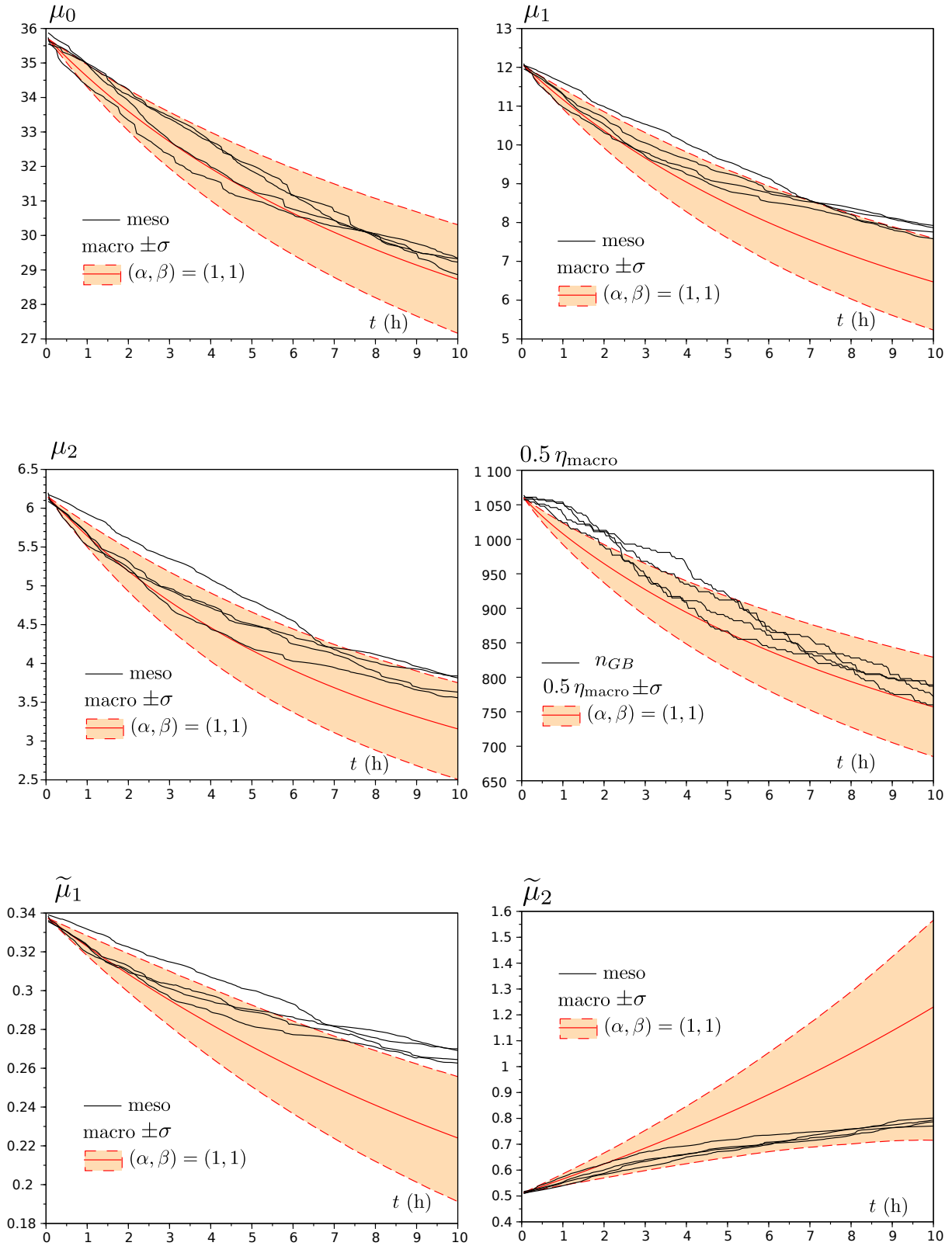


Figure 23: Comparison between 4 overall mesoscopic states and macroscopic state with point-wise standard deviation  $\pm\sigma$ ,  $i_d = 0$ .

known  $\eta_{\text{macro}}$ . Furthermore, the following analysis enables to identify the average grain size evolution. The dimensionless average grain size reads:

$$S_{\text{meso}} = \frac{1}{n} \sum_{k=1}^n S_k \quad (78)$$

where  $n$  is the number of grains in the RVE and  $S_k$  are the dimensionless grain surfaces. A macroscopic estimation of the number of grains  $n_{\text{macro}}$  is obtained from  $\eta_{\text{macro}}$ . Indeed, since for each grain that disappears  $0.5 \dot{\eta}_{\text{macro}} \approx \dot{n}_{GB}$ , and since each grain that disappears is likely to be triangular:

$$\dot{n}_{\text{macro}} \approx \frac{\dot{n}_{GB}}{3} \approx \frac{\dot{\eta}_{\text{macro}}}{6} \quad (79)$$

In addition, the surface of the RVE denoted by  $S = \sum_{k=1}^n S_k$  is constant as detailed in [2] due to mass conservation, that is to say  $\dot{S} = 0$ , thus:

$$\dot{S}_{\text{meso}} = -\frac{\dot{n}}{n^2} S \quad (80)$$

Hence the evolution law of the macroscopic average grain size  $S_{\text{macro}}$  is obtained by replacing in (80) the constant surface of the RVE  $S$  by  $S_{\text{macro}}(t=0)n_{\text{macro}}(t=0)$ :

$$\dot{S}_{\text{macro}} = -\frac{\dot{n}_{\text{macro}}}{n_{\text{macro}}^2} (S_{\text{macro}}(t=0)n_{\text{macro}}(t=0)) \quad (81)$$

where  $n_{\text{macro}}$  is computed from (79). In figure 24 we present the comparison between dimensionless mesoscopic and macroscopic average grain sizes obtained from (78) and (81) respectively. As for the other statistical estimators of the polycrystalline structure, the average size is satisfying as it lies for the most part in the zone defined by more or less one standard deviation.

## 10. Conclusion

In this paper, a macroscopic model of grain growth has been derived as the final step of a general upscaling strategy, which enables to consider directly at the macroscopic scale statistical information of the grain structure, such as grain boundary length, mean and standard deviation of the boundary misorientation distribution etc. The chosen macroscopic variables fully account for the total energy in a deterministic way, which is essential to compute the macroscopic driving force. In addition, the evolution laws of these macroscopic variables have been established in a rigorous way from the mesoscopic evolution law derived in a previous contribution [2]. In the macroscopic evolution laws, a mobility tensor and a vector arise, and have been identified in a *static* approach by using a large database of different and unrelated mesoscopic states in opposition to a *dynamic* approach that would have included mesoscopic evolutions (i.e., successive mesoscopic states). The analysis has shown that an epistemic uncertainty arises from the loss of information due to the reduction of the amount of data encapsulated in the macroscopic state variables in comparison to the mesoscopic state variables. This epistemic uncertainty has been modeled by random variables, whose probability density functions have been estimated by standard Bayesian inference. The macroscopic model is therefore probabilistic and the main result is the mean evolution of the state variables along with point-wise standard deviation.



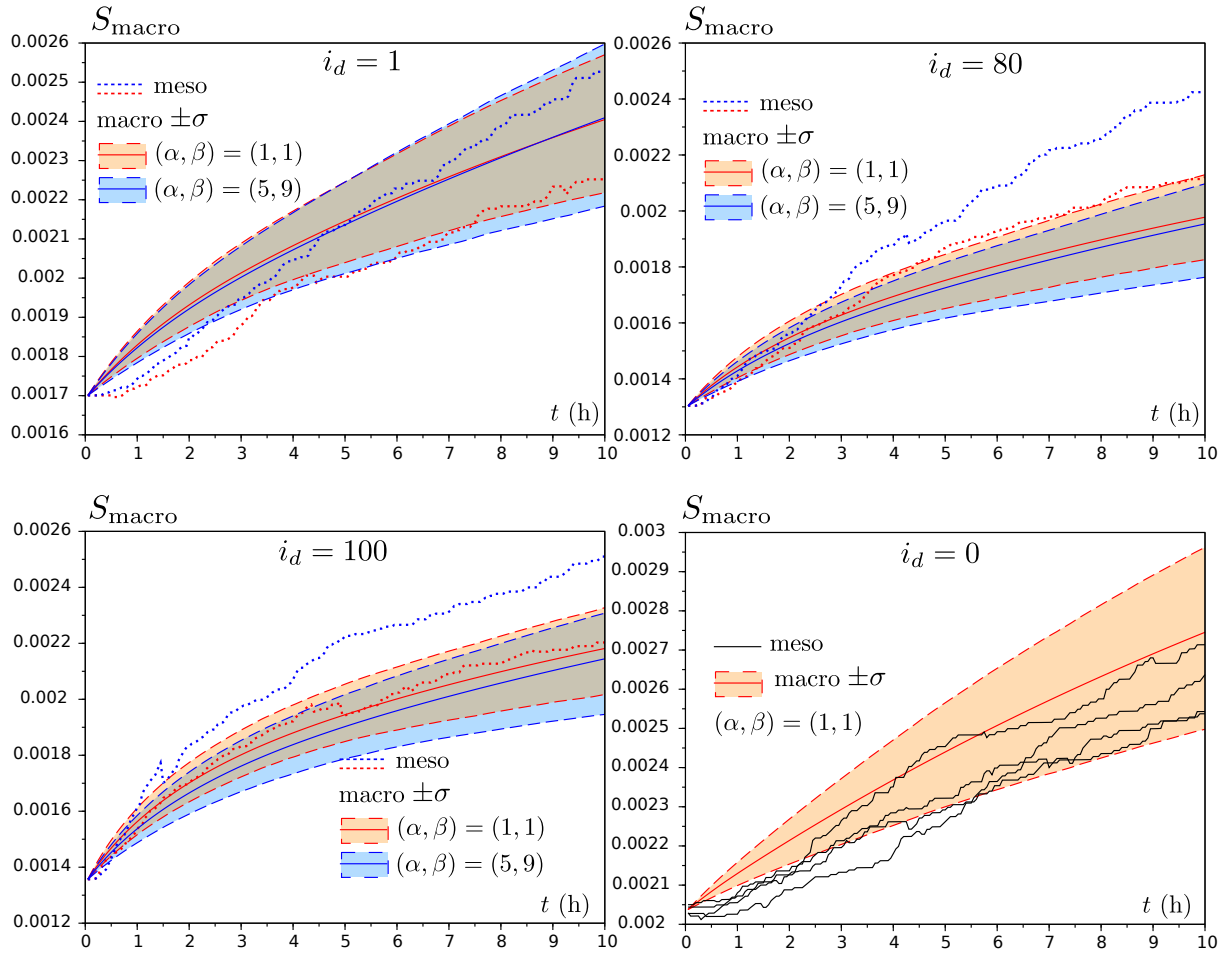


Figure 24: Comparison between dimensionless mesoscopic and macroscopic average grain size with point-wise standard deviation  $\pm\sigma$ .

Results show that the macroscopic model is satisfying especially for the first few hours as good agreement is observed in the comparisons with particular overall states arising from mesoscopic evolutions. Indeed, the mesoscopic evolutions lie for the most part in the zone defined between more or less one standard deviation. In addition, the computation time is fairly reduced as a few seconds are necessary to compute the entire probabilistic macroscopic evolution, which involves the computation of  $N_{\text{stat}} = 500$  particular evolutions. Thus, the proposed approach can be used to carry out simulations of large processes with heterogeneous temperature fields.

The macroscopic model enables to quantify the grain boundary length per unit area  $\mu_0/L_0$  and other statistical descriptors of the polycrystalline structure such as the mean misorientation, its standard deviation or the average grain size. Additional morphological descriptors such as grain circularity have not been included in this study, but there is no fundamental difficulties to adapt the proposed work and include such details.

## References

- [1] J. Humphreys, G. Rohrer, A. Rollett, Recrystallization and related annealing phenomena, Elsevier, 2017. Third edition.

- [2] S. Sakout, D. Weisz-Patrault, A. Ehrlicher, Energetic upscaling strategy for grain growth. I: Fast mesoscopic model based on dissipation, *Acta Materialia* 196 (2020) 261–279.
- [3] D. Weisz-Patrault, S. Sakout, A. Ehrlicher, Fast simulation of grain growth based on orientated tessellation updating method, *Mechanics & Industry* (2020 (in press)).
- [4] K. Hackl, Generalized standard media and variational principles in classical and finite strain elastoplasticity, *Journal of the Mechanics and Physics of Solids* 45 (1997) 667–688.
- [5] D. Kandel, E. Domany, Rigorous derivation of domain growth kinetics without conservation laws, *Journal of Statistical Physics* 58 (1990) 685–706.
- [6] E. A. Holm, G. N. Hassold, M. A. Miodownik, On misorientation distribution evolution during anisotropic grain growth, *Acta Materialia* 49 (2001) 2981–2991.
- [7] E. A. Holm, M. A. Miodownik, A. D. Rollett, On abnormal subgrain growth and the origin of recrystallization nuclei, *Acta Materialia* 51 (2003) 2701–2716.
- [8] J. Gruber, H. Miller, T. Hoffmann, G. Rohrer, A. Rollett, Misorientation texture development during grain growth. part i: Simulation and experiment, *Acta Materialia* 57 (2009) 6102–6112.
- [9] L. Zhang, A. D. Rollett, T. Bartel, D. Wu, M. T. Lusk, A calibrated Monte Carlo approach to quantify the impacts of misorientation and different driving forces on texture development, *Acta Materialia* 60 (2012) 1201–1210.
- [10] A. Kuprat, D. George, G. Straub, M. C. Demirel, Modeling microstructure evolution in three dimensions with grain3d and lagrit, *Computational Materials Science* 28 (2003) 199–208.
- [11] J. Gruber, D. C. George, A. P. Kuprat, G. S. Rohrer, A. D. Rollett, Effect of anisotropic grain boundary properties on grain boundary plane distributions during grain growth, *Scripta materialia* 53 (2005) 351–355.
- [12] H. a. Hallberg, Influence of anisotropic grain boundary properties on the evolution of grain boundary character distribution during grain growth a 2d level set study, *Modelling and Simulation in Materials Science and Engineering* 22 (2014) 085005.
- [13] B. Scholtes, R. Boulais-Sinou, A. Settefrati, D. P. Muoz, I. Poitault, A. Montouchet, N. Bozzolo, M. Bernacki, 3d level set modeling of static recrystallization considering stored energy fields, *Computational Materials Science* 122 (2016) 57–71.
- [14] J. Fausty, N. Bozzolo, M. Bernacki, A 2d level set finite element grain coarsening study with heterogeneous grain boundary energies, *Applied Mathematical Modelling* 78 (2020) 505–518.
- [15] N. Ma, A. Kazaryan, S. Dregia, Y. Wang, Computer simulation of texture evolution during grain growth: effect of boundary properties and initial microstructure, *Acta Materialia* 52 (2004) 3869–3879.

- [16] C. Krill III, L.-Q. Chen, Computer simulation of 3-D grain growth using a phase-field model, *Acta Materialia* 50 (2002) 3059–3075.
- [17] L. Vanherpe, N. Moelans, B. Blanpain, S. Vandewalle, Bounding box framework for efficient phase field simulation of grain growth in anisotropic systems, *Computational Materials Science* 50 (2011) 2221–2231.
- [18] G. Abrivard, E. P. Busso, S. Forest, B. Appolaire, Phase field modelling of grain boundary motion driven by curvature and stored energy gradients. part i: theory and numerical implementation, *Philosophical Magazine* 92 (2012) 3618–3642.
- [19] G. Abrivard, E. P. Busso, S. Forest, B. Appolaire, Phase field modelling of grain boundary motion driven by curvature and stored energy gradients. part ii: application to recrystallisation, *Philosophical Magazine* 92 (2012) 3643–3664.
- [20] K. Chang, N. Moelans, Effect of grain boundary energy anisotropy on highly textured grain structures studied by phase-field simulations, *Acta Materialia* 64 (2014) 443–454.
- [21] M. Upmanyu, D. Srolovitz, L. Shvindlerman, G. Gottstein, Misorientation dependence of intrinsic grain boundary mobility: simulation and experiment, *Acta Materialia* 47 (1999) 3901–3914.
- [22] M. Upmanyu, D. Srolovitz, L. Shvindlerman, G. Gottstein, Molecular dynamics simulation of triple junction migration, *Acta Materialia* 50 (2002) 1405–1420.
- [23] M. Upmanyu, D. J. Srolovitz, A. Lobkovsky, J. A. Warren, W. Carter, Simultaneous grain boundary migration and grain rotation, *Acta Materialia* 54 (2006) 1707–1719.
- [24] F. Humphreys, Modelling mechanisms and microstructures of recrystallisation, *Materials Science and Technology* 8 (1992) 135–144.
- [25] F. Wakai, N. Enomoto, H. Ogawa, Three-dimensional microstructural evolution in ideal grain growth general statistics, *Acta Materialia* 48 (2000) 1297–1311.
- [26] M. Syha, D. Weygand, A generalized vertex dynamics model for grain growth in three dimensions, *Modelling and Simulation in Materials Science and Engineering* 18 (2009) 015010.
- [27] A. Vondrous, M. Reichardt, B. Nestler, Growth rate distributions for regular two-dimensional grains with Read-Shockley grain boundary energy, *Modelling and Simulation in Materials Science and Engineering* 22 (2014) 025014.
- [28] R. Quey, P. Dawson, F. Barbe, Large-scale 3d random polycrystals for the finite element method: Generation, meshing and remeshing, *Computer Methods in Applied Mechanics and Engineering* 200 (2011) 1729–1745.
- [29] H. Telley, T. M. Liebling, A. Mocellin, F. Righetti, Simulating and modelling grain growth as the motion of a weighted voronoi diagram, in: *Materials Science Forum*, volume 94, Trans Tech Publ, pp. 301–306.

- [30] E. Schüle, A justification of the hillert distribution by spatial grain growth simulation performed by modifications of laguerre tessellations, *Computational materials science* 5 (1996) 277–285.
- [31] J. Burke, Some factors affecting the rate of grain growth in metals, *Aime Trans* 180 (1949) 73–91.
- [32] J. Burke, D. Turnbull, *Prog. metal phys.*(eds) b chalmers and r king, 1952.
- [33] M. Anderson, D. Srolovitz, G. Grest, P. Sahni, Computer simulation of grain growth kinetics, *Acta metallurgica* 32 (1984) 783–791.
- [34] J. Von Neumann, *Metal interfaces*, American Society for Metals, Cleveland 108 (1952).
- [35] W. W. Mullins, Two-dimensional motion of idealized grain boundaries, *Journal of Applied Physics* 27 (1956) 900–904.
- [36] D. Moldovan, D. Wolf, S. Phillpot, A. Haslam, Mesoscopic simulation of two-dimensional grain growth with anisotropic grain-boundary properties, *Philosophical Magazine A* 82 (2002) 1271–1297.
- [37] M. Hillert, On the theory of normal and abnormal grain growth, *Acta metallurgica* 13 (1965) 227–238.
- [38] F. Humphreys, A unified theory of recovery, recrystallization and grain growth, based on the stability and growth of cellular microstructures I. the basic model, *Acta Materialia* 45 (1997) 4231–4240.
- [39] A. Rollett, W. Mullins, On the growth of abnormal grains, *Scripta materialia* 36 (1997).
- [40] P. Rios, Comparison between a computer simulated and an analytical grain size distribution, *Scripta materialia* 40 (1999) 665–668.
- [41] W. Fayad, C. Thompson, H. Frost, Steady-state grain-size distributions resulting from grain growth in two dimensions, *Scripta materialia* 40 (1999) 1199–1204.
- [42] F. Fischer, J. Svoboda, P. Fratzl, A thermodynamic approach to grain growth and coarsening, *Philosophical Magazine* 83 (2003) 1075–1093.
- [43] J. Svoboda, F. Fischer, A new approach to modelling of non-steady grain growth, *Acta materialia* 55 (2007) 4467–4474.
- [44] F. Fischer, J. Svoboda, H. Petryk, Thermodynamic extremal principles for irreversible processes in materials science, *Acta materialia* 67 (2014) 1–20.
- [45] D. Piot, G. Smagghe, J. J. Jonas, C. Desrayaud, F. Montheillet, G. Perrin, A. Montouchet, G. Kermouche, A semitopological mean-field model of discontinuous dynamic recrystallization, *Journal of materials science* 53 (2018) 8554–8566.
- [46] P. De Micheli, L. Maire, D. Cardinaux, C. Moussa, N. Bozzolo, M. Bernacki, Digimu®: Full field recrystallization simulations for optimization of multi-pass processes, in: *AIP Conference Proceedings*, volume 2113, AIP Publishing LLC, p. 040014.

- [47] W. T. Read, W. Shockley, Dislocation models of crystal grain boundaries, *Physical review* 78 (1950) 275.
- [48] D. Wolf, A Read-Shockley model for high-angle grain boundaries, *Scripta Metallurgica* 23 (1989) 1713–1718.
- [49] M. D. Hoffman, A. Gelman, The no-u-turn sampler: adaptively setting path lengths in hamiltonian monte carlo., *Journal of Machine Learning Research* 15 (2014) 1593–1623.
- [50] J. Salvatier, T. V. Wiecki, C. Fonnesbeck, Probabilistic programming in python using pymc3, *PeerJ Computer Science* 2 (2016) e55.
- [51] G. Van Rossum, F. L. Drake, *The python language reference manual*, Network Theory Ltd., 2011.
- [52] R. G. Ghanem, P. D. Spanos, *Stochastic finite elements: a spectral approach*, Courier Corporation, 2003.