Bayesian inference of grain growth prediction via multi-phase-field models

Shin-ichi Ito, ¹ Hiromichi Nagao, ^{1,2} Takashi Kurokawa, ^{2,*} Tadashi Kasuya, ³ and Junya Inoue^{4,3} ¹Earthquake Research Institute, The University of Tokyo, 1-1-1, Yayoi, Bunkyo-ku, Tokyo 113-0032, Japan ²Graduate School of Information Science and Technology, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan ³Graduate School of Engineering, The University of Tokyo, 7-3-1, Hongo, Bunkyo-ku, Tokyo 113-8656, Japan ⁴Research Center for Advanced Science and Technology, The University of Tokyo, 4-6-1, Komaba, Meguro-ku, Tokyo 153-8904, Japan



(Received 18 September 2018; published 30 May 2019)

We propose a Bayesian inference methodology to evaluate unobservable parameters involved in multi-phasefield models to accurately predict the observed grain growth, such as in metals and alloys. This approach integrates models and a set of observational image data of grain structures. Because the image data set is not a time series, directly applying conventional inference techniques that require time series as the input data is difficult. The key idea in our methodology to overcome this difficulty is to construct a time series with an appropriate statistic that characterizes static image data of grain structures. Our methodology implements the empirical Bayes method. It can estimate not only a probability density function of the parameters but also an initial phase field, which is generally unobservable in real experiments. After validating the proposed method through numerical tests using synthetic data, we apply it to real experimental images of grain structures in a steel alloy. The proposed method properly estimates unobservable parameters along with their uncertainties and successfully selects the initial phase field that best explains the experimental data from among candidate initial phase fields.

DOI: 10.1103/PhysRevMaterials.3.053404

I. INTRODUCTION

Controlling the mechanical properties of structural materials such as iron, stainless steel, and aluminum is a crucial issue in materials science. A fundamental factor that determines structural properties is the microstructure, which is composed of grains with different crystallographic orientations. Predicting the temporal grain evolution is a key task for efficiently developing materials, because it is directly related to control of the macroscopic mechanical properties [1-3]. The time evolution of the grain structure is generally driven by interactions between neighboring grains due to the curvatures of grain boundaries [4,5], diffusion of chemical elements [6,7], and thermal effects [8,9]. In prediction, numerical simulations that consider these interactions are essential for controlling the dynamics of the grain structure needed to develop materials with the desired mechanical properties. The phase-field (PF) method is widely accepted for simulating the dynamics of grain structures [10–12]. The key parameters that determine the dynamics of grain growth are interfacial energy γ , PF mobility L, and thickness W of grain boundaries. Although accurately determining these parameters and the initial PF needed to conduct a PF simulation is essential for the accurate prediction of the growth of the grain structure, these prerequisites are typically assumed based on experience or toy experiments, which may cause large prediction errors. However, the recent innovation [13–17] of statistical methodologies based on Bayesian inference (BI) [18] enables us to identify them by comparing the results of PF simulations and limited observational data.

The BI method evaluates a probability density function (PDF) (precisely, a posterior PDF) of unmeasurable or unobservable parameters involved in given simulation models by integrating the simulation models and limited observational or experimental data. The posterior PDF contains rich information, e.g., the optimum and its uncertainty, given by the maximum and broadness of the PDF, respectively, which is needed to update the model and observational or experimental designs. In particular, data assimilation (DA) [16], which is a kind of BI mainly developed in the fields of the meteorology and oceanography [19–21], constructs the posterior PDF based on time series of data and forecasts based on the PDF [13–15]. Because the integration of physical models and observational or experimental data is an essential procedure in various scientific areas, DA has been expanding into application fields such as seismology [22], biology [23], petrology [24], and materials science [17,25-27]. DA, which can be directly applied when an observational dataset is given as a time series, systematically estimates parameters that determine the dynamics of a given system, e.g., γ , L, and W in the PF models. However, the time series of data is not always available in materials science. Therefore, in constructing a DA methodology for materials science, we should bear in mind the following: (i) observational or experimental data are often static images of grain structures and (ii) ordinary DA requires us to compare the time series of a PF model and the time series of data, including information about the time evolution of grain structures.

This study proposes a BI-based methodology to estimate γ , L, and W by designing appropriate time series of a statistic

^{*}Present address: Yahoo Japan Corporation, Japan.

that characterizes the grain structure in each static image. Our methodology relies on an empirical Bayes method (EBM) that simultaneously estimates the parameters along with their uncertainties and the initial PF, which is generally unobservable in real experiments. We apply the proposed method to real experimental data of austenite grain structures, which exhibit purely curvature-driven grain-boundary growth.

The remainder of this paper is organized as follows: Section II introduces the PF model used in this study and then discusses the significance of estimating the parameters involved in the model from the perspective of application to real data. Section III proposes an appropriate statistic for practical image data of grain structures, which is to be input into the BI procedure mentioned in Sec. IV. The proposed methodology is validated through a numerical test using synthetic data in Sec. V. An application to real data and discussions are given in Sec. VI, and we conclude in Sec. VII.

II. MULTI-PHASE-FIELD MODEL

This study describes the temporal evolution of grain structures using a multi-phase-field (MPF) model [10,11]. The key point in this model is that grains are described by a set of time-dependent field variables $\{\phi_i(\mathbf{x},t)\}\$, termed PF variables, where x and t denote place and time, respectively. Let N be the number of grains (or PF variables) that exist at the start of the simulation. The PF variable $\phi_i(x, t)$ indicates the existence probability of the grain i (i = 1, ..., N), which satisfies the conservation law $\sum_{i=1}^{N} \phi_i(\mathbf{x}, t) = 1$. The condition $\phi_i(\mathbf{x}, t) =$ 1 implies that only grain i occupies the position x, $\phi_i(x, t) = 0$ implies that grain i does not exist at x, and $0 < \phi_i(x, t) < 1$ implies that the grain boundaries between grain i and the other grains exist at x. Considering the physical properties of the desired material mentioned in Sec. III A, the MPF model used in this study assumes that the dynamics of the grain structure are driven only by the curvature effect of the grain boundaries. The MPF model proposed by Steinbach et al. [10], upon which this study is based, is

$$\frac{\partial \phi_i}{\partial t} = -\frac{2}{n} \sum_{k=\pm i}^{N} \sum_{j=\pm k}^{N} s_i s_j (f_{ijk} - f_{jik}), \tag{1}$$

where $s_i(\mathbf{x}, t)$ is an indicator function, i.e., $s_i(\mathbf{x}, t) = 1$ if $0 < \phi_i(\mathbf{x}, t) \le 1$ and $s_i(\mathbf{x}, t) = 0$ if $\phi_i(\mathbf{x}, t) = 0$, and the integer $n(\mathbf{x}, t)$ is the number of grains that exist at \mathbf{x} , i.e., $n = \sum_{i=1}^{N} s_i$. The tensor f_{ijk} that appears in Eq. (1) is given by

$$f_{ijk} = L_{ij} \left(\frac{\epsilon_{ik}^2}{2} \nabla^2 \phi_k + \omega_{ik} \phi_k \right), \tag{2}$$

where L_{ij} is the PF mobility and ϵ_{ij} and ω_{ij} are described by using the interfacial energy γ_{ij} and the thickness W_{ij} of the grain boundary, respectively [9], as

$$\epsilon_{ij} = \frac{2}{\pi} \sqrt{W_{ij} \gamma_{ij}},\tag{3}$$

$$\omega_{ij} = \frac{4\gamma_{ij}}{W_{ii}}. (4)$$

Originally, these parameters are defined for each pair of grains i and j; however, the subscripts can be omitted by replacing

the parameters with their average values, i.e., $L_{ij} = L$, $\epsilon_{ij} = \epsilon$, $\omega_{ij} = \omega$, $W_{ij} = W$, and $\gamma_{ij} = \gamma$, because the crystallographic orientations of the grains are assumed to be distributed randomly and isotropically in this study. In this case, Eq. (1) is reduced to

$$\frac{\partial \phi_i}{\partial t} = s_i \left[A \left(\phi_i - \frac{1}{n} \right) + B \left(\nabla^2 \phi_i - \frac{1}{n} \sum_{j=1}^N s_j \nabla^2 \phi_j \right) \right], \quad (5)$$

where $A=2L\omega$ and $B=L\epsilon^2$ are the free parameters associated significantly with the dynamics, which are to be estimated from observational data. These parameters A and B relate to the three free parameters contained in the MPF model, i.e., γ , L, and W, as follows:

$$W = \pi \sqrt{\frac{2B}{A}},\tag{6}$$

$$\gamma L = \frac{\pi}{4\sqrt{2}}\sqrt{AB}.\tag{7}$$

Note that the interfacial energy γ is a physical parameter that can be measured in real experiments, whereas the thickness of grain boundaries W and the PF mobility L are nonphysical parameters that are difficult to determine directly from real experiments. Conventional studies determine W and L by trial and error, which requires many PF simulations to reproduce results that support experimental results. The actual computational cost needed to conduct a PF simulation is generally high; therefore, determinations by trial and error require unrealistically burdensome computations that often exceed the limitations of available computer resources. Moreover, because the determinations are usually conducted on the basis of experience and intuition, quantifying the underlying uncertainties in W and L is difficult. Because such uncertainties affect the simulation results, we need to develop a methodology to estimate W and L along with their uncertainties through the integration of the PF model and observational data. Furthermore, the comparison between the parameters estimated by this methodology and those obtained by first-principle calculations, such as the molecular dynamics simulations [28,29], enables us to validate and improve the given PF model. Thus, establishing a methodology to evaluate W and L is significant from the perspective of PF simulation studies.

III. EXPERIMENTS

A. Target material

The target material in this study is Fe-0.27C-0.18Si-0.45Mn-0.014P-0.003S steel alloy, in which the grain structure evolves with time in an isothermal environment. Because this steel does not include elements that may hinder austenite grain growth, we can observe purely curvature-driven grain growth. In addition, the experiments show that the crystallographic orientations of the grains in this steel are randomly distributed, which ensures that the MPF model shown by Eq. (5) can describe the grain-growth dynamics. After setting a sample of the steel into a heating furnace, we conduct a heating treatment involving the following three stages: heating, holding, and cooling, as

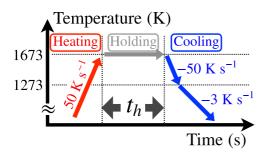


FIG. 1. Heat treatment protocol for obtaining a sample of the grain structure with holding time t_h .

described in Fig. 1. In the heating stage, the temperature in the furnace increases to 1637 K starting from room temperature with a high heating rate of 50 K s^{-1} . When the temperature reaches 1637 K, heating stops, and the holding stage starts. In the holding stage, the temperature in the furnace is maintained at 1637 K. The characteristic size of evolving grains depends on the holding time t_h , which indicates how long the sample is kept in this isothermal environment. After t_h has elapsed, the cooling stage starts to stop the growth of grain structures. In this stage, the temperature of the furnace decreases to 1237 K with a high cooling rate of -50 K s^{-1} to approach room temperature at -3 K s^{-1} . The rapid cooling promotes ferrite nucleation at the austenite grain boundary, which simplifies boundary detection.

This study applies this heat treatment to five samples of steel using different holding times t_h to obtain grain structures with different characteristic grain sizes. Figure 2 shows the obtained grain structures for the holding times (a) $t_h = 1$ s,

(b) $t_h = 2$ s, (c) $t_h = 5$ s, (d) $t_h = 10$ s, and (e) $t_h = 20$ s, respectively. The red solid lines indicate the grain boundaries, whereas the polygons surrounded by the lines indicate grains. The characteristic grain size is confirmed to increase with increasing holding time t_h ; thus, the series of these images synthetically shows the temporal evolution of grain structure. Note that the direct application of the ordinary DA to this images series is impossible because the series was obtained from different samples.

B. Design of observational data

Based on the images of grain structures shown in Fig. 2, we designed observational data comparable to the results of the MPF simulations. Because the PF parameters to be estimated have the dimensions of time and length, the dimensions of the observational data should also be of length or its higher orders. In this study, we employed a statistic related to the characteristic areal size of grains as the observational data, because this statistic is ensured to be identical on the assumption that the underlying probability density of the grain size depends only on t_h as the number of grains approaches infinity. Each grain size in Fig. 2 can be easily computed by measuring the area of the closed polygon denoting the grain. Let N_{t_h} be the number of grains in the snapshot at t_h and $\{S_1, \ldots, S_{N_{t_h}}\}$ be the set of grain sizes. We propose a statistic "moment ratio" K given by

$$K(t_h) = \frac{\sum_{i=1}^{N_{t_h}} S_i^2}{\sum_{i=1}^{N_{t_h}} S_i}.$$
 (8)

Figure 3 shows the moment ratio K as a function of t_h computed from the grain structures shown in Fig. 2. This statistic has the dimensions of areal size and exhibits

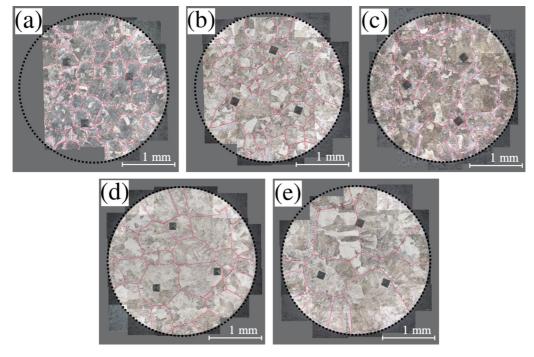


FIG. 2. Grain structures of the target steel obtained by heat treatment with holding times (a) $t_h = 1$ s, (b) $t_h = 2$ s, (c) $t_h = 5$ s, (d) $t_h = 10$ s, and (e) $t_h = 20$ s. The dotted black circles 3 mm in diameter indicate the surface domains of the samples. The red solid lines indicate the grain boundaries.

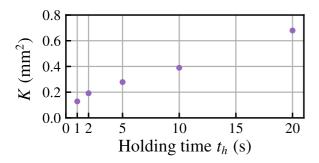


FIG. 3. Moment ratio K generated by the grain structures shown in Fig. 2 as a function of the holding time t_h .

similar behavior to that of the average grain size predicted by Hillert's mean-field analysis of curvature-driven grain growth [4,30]. This implies that the moment ratio K properly captures the temporal evolution of the characteristic grain size in curvature-driven grain growth.

We emphasize here the advantages of using this moment ratio K as the observational data rather than other statistics such as the average grain size used previously. Because ordinary DA requires comparing the time series of observational data and simulation results, it is ideal to use the time series of an identical grain structure obtained by $in \ situ$ observation in an isothermal furnace. However, such $in \ situ$ observation is technically difficult owing to the contamination of the steel surface by oxidation. In addition, some treatments such as polishing and etching are usually needed prior to the observation of the fine grain structure, making acquisition of the time series of an identical grain structure difficult. Our proposed statistic enables comparison with the simulation results, because the statistic absorbs the details of the grain structures arising from individual sample differences.

IV. BAYESIAN INFERENCE

This section presents the BI methodology used to estimate the PF parameters A and B and the initial PF $\Phi = \{\phi_i(x, t = 1s)\}$ involved in Eq. (5). The key technique here is EBM [31,32], which enables us to obtain a posterior PDF related to the PF parameters and simultaneously select the best initial PF from among several candidates.

A. Phase-field model quantities for comparison with observations

Let $\mathcal{T} = \{t_1, \dots, t_{n_d}\}$ be the set of t_h sorted in ascending order, where n_d is the size of \mathcal{T} , and let $\mathbf{D} = \{K(t_1), \dots, K(t_{n_d})\}$ be the set of K values obtained from the observations at times in the set \mathcal{T} . To compare observations with results from the MPF simulations, a model-based quantity corresponding to K must be extracted from the PF variables. We define this comparative quantity as a model-based moment ratio $\mathcal{K}(t)$:

$$\mathcal{K}(t) = \frac{\sum_{i=1}^{N} \mathcal{S}_i(t)^2}{\sum_{i=1}^{N} \mathcal{S}_i(t)},$$
(9)

TABLE I. Parameters for MPF simulations.

Parameter	Notation	Value
Number of grains	N	80
Length of a side of the	H	2.7 mm
computational domain		
Grid spacing	Δx	0.021 mm
Time step	Δt	$1.12 \times 10^{-4} \text{ s}$

where $S_i(t)$ is the areal size of grain i, which is computed using the PF ϕ_i as

$$S_i(t) = \int d\mathbf{x} \,\phi_i^{1+\alpha}(\mathbf{x}, t). \tag{10}$$

The positive parameter α tunes the sharpness of the grain boundaries, letting $\alpha = 4$ in this study. In general, $S_i(t)$ approximates the actual areal size of the ith grain more accurately as α tends to infinity. Simple numerical experiments show that the approximation is sufficient when $\alpha \geqslant 4$, but it gets worse when $\alpha < 4$. We demonstrated that this modelbased moment ratio K properly describes the behavior of the data-based moment ratio K by conducting the MPF simulations with various sets of parameters. The temporal evolutions of the MPF model were computed with a two-dimensional square periodic spacing with dimensions of $H \times H$ starting from an initial PF structure at t = 1 s. The initial PF structure was composed of N grains, the locations of which were tuned so that their computed model-based moment ratio K was equal to the data-based moment ratio K at t = 1 s. To solve the MPF model numerically, we discretized the computational domain by using a square lattice with a grid spacing Δx and applied the computational scheme proposed by Kim et al. [12]. This scheme reduces the computational cost to $O(m_{\text{time}}m_{\text{grid}})$, where m_{grid} is the number of grid points and m_{time} is the total number of time steps, which is considerably less than the cases when naïve schemes are implemented. A time step Δt for time discretization was selected to prevent the simulation from diverging. Table I lists the parameters used in the numerical simulations. Figure 4 exemplifies temporal evolutions of the model-based moment ratio \mathcal{K} computed from the solutions of Eq. (5) in the three cases of the parameter A. The computed time series of K and the grain structures seemed roughly consistent with the values of K (Fig. 3) and real observations (Fig. 2).

B. Observation model

Our BI method requires an "observation model" that describes the relation between observations and simulations. This study assumed that the data-based moment ratio K(t) [Eq. (8)] was equivalent to the model-based moment ratio $\mathcal{K}(t)$ [Eq. (9)] contaminated by noise, i.e.,

$$K(t_i) = \mathcal{K}(t_i) + \Omega_i \quad (t_i \in \mathcal{T}),$$
 (11)

where the observation noise Ω_i ($i = 1, ..., n_d$) was assumed to identically and independently follow a normal distribution $q(\Omega)$ with a mean of zero and a constant variance of σ^2 , i.e.,

$$q(\Omega) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\Omega^2}{2\sigma^2}\right). \tag{12}$$

Here, we assume that σ is an undetermined hyperparameter.

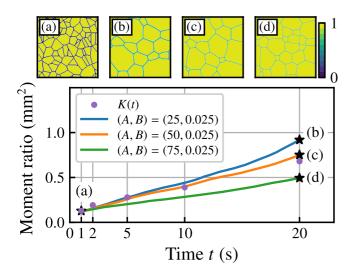


FIG. 4. Data-based moment ratio K(t) (the same as in Fig. 3) and model-based moment ratios with three sets of A and B (bottom), starting from an identical initial grain structure at t=1 s [top (a)]. The grain structures and corresponding moment ratios at t=20 s are indicated as (b)–(d), where the colors in the top panels are the magnitudes of $\sum_{i=1}^{N} \phi_i^2$. The units of A and B are s^{-1} and $mm^2 s^{-1}$, respectively.

C. Bayes' theorem

Here, we construct a posterior PDF, which is a conditional PDF of the parameters A and B, and an initial PF Φ' , conditional on a given dataset D, based on Eqs. (11) and (12). Bayes' theorem states that a posterior PDF $p(A, B, \Phi'|D, \sigma)$ is given by

$$p(A, B, \Phi' | \mathbf{D}, \sigma) \propto p(A, B) p(\Phi') p(\mathbf{D} | A, B, \sigma, \Phi'),$$
 (13)

where p(A, B) is the joint prior PDF of the parameters A and B, $p(\Phi')$ is the prior PDF of the initial PF Φ' , and $p(D|A, B, \sigma, \Phi')$ is a likelihood function. This study assumed that the initial PF Φ' is not stochastic but deterministic, so that the prior PDF $p(\Phi')$ is given as

$$p(\Phi') = \delta(\Phi' - \Phi), \tag{14}$$

where $\delta(\bullet)$ is the Dirac δ function. However, the prior PDF p(A, B) is assumed to be a uniform distribution:

$$p(A, B) = \begin{cases} \frac{1}{A_{\text{max}} B_{\text{max}}} & \text{when} \quad (A, B) \in \Gamma \\ 0 & \text{otherwise,} \end{cases}$$
 (15)

where Γ is a rectangular region given by

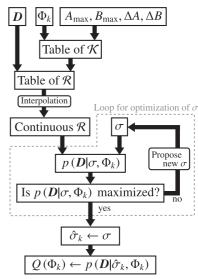
$$\Gamma = \left\{ \begin{pmatrix} A \\ B \end{pmatrix} \in \mathbb{R}^2 \middle| 0 \leqslant A \leqslant A_{\text{max}}, 0 \leqslant B \leqslant B_{\text{max}} \right\}$$
 (16)

The likelihood function $p(\mathbf{D}|A, B, \sigma, \Phi')$ measures the consistency between the data and the model, which is defined here based on $q(\Omega)$ as

$$p(\mathbf{D}|A, B, \sigma, \Phi') = \prod_{t_i \in \mathcal{T}} q(\Omega_i)$$

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^{n_d} \exp\left(-\frac{\mathcal{R}}{2\sigma^2}\right), \quad (17)$$

Prepare candidates $V = {\Phi_1, ..., \Phi_{n'}}$ For k = 1, ..., n'



End for

FIG. 5. Procedure to select the best initial phase field from a candidate set $V = \{\Phi_1, \dots, \Phi_{n'}\}$ based on the score Q.

where \mathcal{R} is the sum of squared residuals, i.e.,

$$\mathcal{R} = \sum_{t_i \in \mathcal{T}} [K(t_i) - \mathcal{K}(t_i)]^2.$$
 (18)

Note that \mathcal{R} depends on A, B, and Φ' through the temporal evolution of the PF resulting from the MPF simulations. Substituting Eqs. (14)–(18) for Eq. (13), we obtain a posterior PDF $p(A, B, \Phi'|D, \sigma)$, but this is not the objective posterior PDF related to the parameters A and B. The objective can be obtained by marginalizing Eq. (13) with respect to Φ' as

$$p(A, B|\mathbf{D}, \sigma, \Phi) = \frac{p(A, B)p(\mathbf{D}|A, B, \sigma, \Phi)}{p(\mathbf{D}|\sigma, \Phi)}, \quad (19)$$

where the normalization constant $p(\mathbf{D}|\sigma, \Phi)$ is given by

$$p(\mathbf{D}|\sigma,\Phi) = \int_{-\infty}^{\infty} dA \int_{-\infty}^{\infty} dB p(A,B) p(\mathbf{D}|A,B,\sigma,\Phi).$$
 (20)

To fully determine the posterior PDF of Eq. (19), the unknown σ , which is unmeasurable directly from K(t), must be given. Thus, this study employed an EBM as follows:to determine σ objectively.

D. Empirical Bayes method

Our EBM-based methodology is capable not only of determining σ objectively but also selecting the best initial PF from a given candidate set (Fig. 5) as follows:

The EBM [31,32], which is one of the BI methods, determines hyperparameters involved in a given statistical model based on the maximum likelihood principle. Let p(Y|Z) be a PDF of observational data Y, where Z denotes a hyperparameter vector that characterizes the PDF. We aim to estimate an optimum hyperparameter vector \hat{Z} for Z by using realizations Y^* through evaluating the goodness of the PDF. When Y^* is given, the goodness is given by the likelihood function

 $p(Y^*|Z)$, i.e., the optimum \hat{Z} is obtained by maximizing the likelihood function with respect to Z. In our case, Eq. (20) corresponds to the PDF to be maximized with given data D, so that an optimum $\hat{\sigma}$ for σ can be obtained as

$$\hat{\sigma} = \underset{\sigma}{\operatorname{argmax}} p(\mathbf{D}|\sigma, \Phi)$$

$$= \underset{\sigma}{\operatorname{argmax}} \int_{-\infty}^{\infty} dA \int_{-\infty}^{\infty} dB p(A, B) p(\mathbf{D}|A, B, \sigma, \Phi)$$

$$= \underset{\sigma}{\operatorname{argmax}} \int_{0}^{A_{\text{max}}} dA \int_{0}^{B_{\text{max}}} dB p(\mathbf{D}|A, B, \sigma, \Phi). \tag{21}$$

This type of methodology for estimating hyperparameters, which needs to optimize the normalization constant, is termed the EBM. As seen in Eq. (21), the naïve implementation of the EBM incurs a high computational cost because a large number of likelihood function calculations are required. Hence, the EBM is rarely performed in massive simulation models such as PF models. However, in our case, the EBM is available with low computational cost because the simultaneous computation of R and σ is not required when computing the likelihood $p(D|A, B, \sigma, \Phi)$. In practical computation, prior to evaluating the integral, we prepare a table of values of the model-based moment ratio K that are evaluated at discrete points in Γ . The table can be utilized to compute a table of \mathcal{R} at the discrete points; subsequently, a continuous field of \mathcal{R} is reproduced by using the table of \mathcal{R} via interpolation. The continuous field can be reproduced with a low computational cost; hence, we can perform the EBM even when with a massive simulation model. In this study, we discretize Γ onto a rectangular mesh with the grid spacings ΔA for A and ΔB for B and adopt a first-order interpolation to reproduce the continuous \mathcal{R} . By substituting the obtained $\hat{\sigma}$ in Eq. (19), we obtain the full form of the posterior PDF $p(A, B|D, \hat{\sigma}, \Phi)$. Furthermore, our methodology is available for the selection of the optimum initial PF from several candidates assumed in advance. Let $V = \{\Phi_1, \dots, \Phi_{n'}\}$ be a set of candidates for the initial PFs, where n' is the number of candidates, and let $\hat{\sigma}_k$ be the $\hat{\sigma}$ obtained by applying our EBM-based methodology to an initial PF Φ_k . From the perspective of the EBM, the selection is performed by evaluating a score Q given by

$$Q(\Phi_k) = p(\mathbf{D}|\hat{\sigma}_k, \Phi_k) \quad (k = 1, \dots, n'), \tag{22}$$

which is naturally obtained when computing $\hat{\sigma}_k$. The initial PF with the largest score is selected as the best one. Note that a simple fitting of the data-based moment ratio K does not provide information related to the initial grain structure, whereas our EBM-based methodology can simultaneously estimate the parameters and select the best initial grain structure.

V. TWIN EXPERIMENTS

We conducted "twin experiments" that validated the proposed method based on synthetic data generated by the simulation model using true parameters assumed in advance [33]. The true parameters were assumed to be $A_{\rm true} = 200 \, {\rm s}^{-1}$ for A, and $B_{\rm true} = 0.075 \, {\rm mm}^2 \, {\rm s}^{-1}$ for B. The simulation parameters listed in Table I were used in all twin experiments. A MPF simulation using these parameters generated synthetic obser-

vational data K at t = 1, 2, 5, 10, and 20 s, following which synthetic observation noise following a normal distribution with a mean of zero and a constant variance of σ^2 contaminated the data. This study conducted the following three twin experiments: Twin experiment I (Sec. VA) investigated the influence of the settings of the domain Γ on estimation results. As shown in Eq. (21), the upper bounds A_{max} and B_{max} obtained from experience and intuition are directly related to the accuracy of the posterior PDF. Twin experiment II (Sec. VB) investigated the influence of signal-to-noise ratios for the observations, which depends on the magnitude of the standard deviation σ . The accuracy of estimation of σ directly influences the reliability of the posterior PDF of A and B. Twin experiment III investigated the selection of the best initial PF from among the candidates. The score Q [Eq. (22)] was shown to be an appropriate indicator for the selection of the initial PF that best matches the synthetic observational data.

A. Twin experiment I: Influence of the domain Γ

The setting of the domain Γ is an important factor in obtaining a proper estimation. A wider Γ would give a more objective estimation result; however, it usually requires a massive computational cost. Twin experiment I aims at finding a minimum area for Γ that provides a correct estimation. In general, successful estimation is not always ensured when using a newly proposed estimation method incorporating subjective settings, such as Γ in our case. Hence, we checked the validity of our method by investigating the dependence of the estimation results on Γ . To perform our methodology, Γ is discretized with grid spacings of $\Delta A = 6.25 \text{ s}^{-1}$ for A and $\Delta B = 0.001\,562\,5\,\text{ mm}^2\,\text{s}^{-1}$ for B. For generating synthetic observational data, the initial PF was assumed to be the same as the one used in Sec. IV A, and the standard deviation σ of the observation noise used to contaminate the synthetic observational data was set to be $\sigma = 10^{-5}$ mm². Whether the maximizer, which is a set of A and B that maximizes the posterior PDF, is consistent with the true parameters assumed in advance is of interest in this experiment. We performed the proposed method many times, by changing the values of A_{max} and B_{max} , i.e., changing the area of the domain Γ [Fig. 6(a)]. Figures 6(b) and 6(c) show the maximizer (\hat{A}, \hat{B}) for various sets of $(A_{\text{max}}, B_{\text{max}})$. This indicates that the estimation always succeeds when Γ includes the set $(A_{\text{true}}, B_{\text{true}})$, while the estimation does not work well when Γ does not include the set. Although the latter fact is trivial, the former fact is nontrivial because successful estimation, in general, is not always ensured, even if Γ includes the true parameters. Our methodology ensures estimation of the true parameters when Γ includes them; therefore, we conclude that our methodology provides an essential ability as a parameter estimator.

B. Twin experiment II: Influence of signal-to-noise ratio

The signal-to-noise ratio is also a factor that affects the quality of the posterior PDF. Twin experiment II investigated the influence of this factor in the case of Gaussian noise with varied noise levels, i.e., the variance σ^2 . The bounds $A_{\text{max}} =$

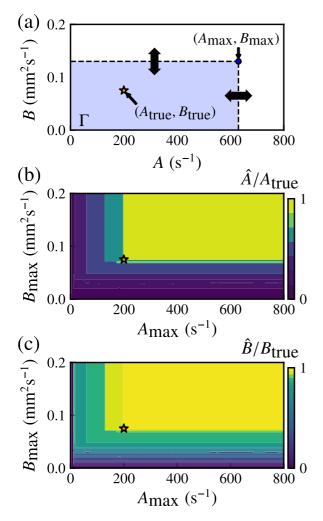


FIG. 6. (a) Setup of Twin experiment I. The blue area indicates the domain Γ defined by Eq. (16), which determines the prior for A and B. Twin experiment I investigated the influence of Γ in the estimation of the maximizer (\hat{A}, \hat{B}) . The star indicates the true parameter set. (b), (c) Results of Twin experiment I. The colors, the values of \hat{A}/A_{true} and \hat{B}/B_{true} , indicate how appropriately the maximizer (\hat{A}, \hat{B}) was estimated when the domain Γ was varied. The stars denote the true case, i.e., $(A_{\text{max}}, B_{\text{max}}) = (A_{\text{true}}, B_{\text{true}})$.

800 s⁻¹ and $B_{\rm max}=0.2~{\rm mm^2~s^{-1}}$ were fixed, and the other settings except for σ were the same as in Twin experiment I. Figure 7 shows the average and the 95th and 5th percentiles of the estimated $\hat{\sigma}$ obtained from a thousand trials with different observation noise. The averages lie clearly on the line $\hat{\sigma}=\sigma$, which indicates the validity of the proposed method.

C. Twin experiment III: Selection of initial phase fields

Twin experiment III confirmed that the score Q is an appropriate indicator for selecting an initial PF from among the given candidates. Figure 8(a) shows the three candidate initial PFs $V = \{\Phi_1, \Phi_2, \Phi_3\}$ for this experiment, which were generated by the procedure described in Sec. IV A. Each of these candidates satisfies the condition that the initial value of \mathcal{K} is equal to the value of the data-based moment ratio K at t = 1 s. The synthetic observational data $\mathcal{K}(t)$ were computed

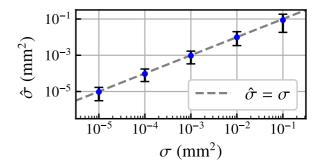


FIG. 7. Results of Twin experiment II. The relation between the estimated $\hat{\sigma}$ and assumed σ is shown. For each σ , a thousand trials of estimations of $\hat{\sigma}$ with different observation noise were conducted. Each blue point indicates the average, and the upper and lower ticks in each error bar indicate the 95th and 5th percentiles, respectively. The dashed gray line indicates $\hat{\sigma} = \sigma$.

by the MPF model by using the parameters in Table I and an initial PF chosen from V (we call this the "true initial PF"). Then, the data were contaminated by observational noise that followed a normal distribution with a mean of zero and a standard deviation of $\sigma = 0.001$ mm². Using this synthetic observational data, we computed the score for each candidate based on Eq. (22). If the initial PF with the highest score out of the three candidates was consistent with the true initial PF, the proposed method was validated. Figure 8(b) shows three experimental results of the scores in cases where the true initial PF was given by Φ_1 , Φ_2 , and Φ_3 . In each experiment, the score of an initial PF was largest when it was consistent with the true initial PF. This indicates the validity of the score Q as an indicator for the selection of the initial PF.

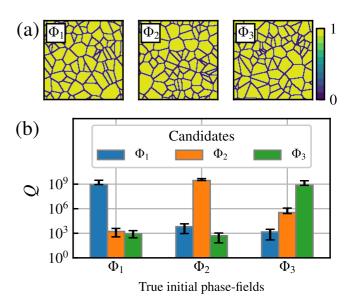


FIG. 8. Results of Twin experiment III. (a) Initial phase fields used in this experiment. (b) Comparison of the scores Q. The horizontal axis indicates the true initial phase fields, i.e., the initial phase fields used to generate the synthetic observational data. The height of each bar indicates the average of Q values over a thousand trials with different observational noise. The upper and lower ticks in each error bar indicate the 95th and 5th percentiles, respectively.

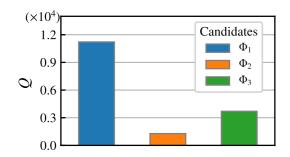


FIG. 9. Selection of the initial phase field based on the score Q in the case of real observational data. The candidate initial phase fields are shown in Fig. 8(a).

VI. APPLICATION TO REAL DATA

The proposed methodology was applied to the real databased moment ratio K shown in Fig. 3. The domain Γ was set letting $A_{\text{max}} = 800 \text{ s}^{-1}$ and $B_{\text{max}} = 0.2 \text{ mm}^2 \text{ s}^{-1}$, which sufficiently covered the parameter space to provide robust estimation. The other parameters in the MPF simulations were set as in Table I. The candidate initial PFs were as shown in Fig. 8(a). Figure 9 shows that Φ_1 was the best initial PF from among $\{\Phi_1, \Phi_2, \Phi_3\}$, as indicated by the score Q for Φ_1 being significantly larger than the scores for the other candidates. Figure 10(a) shows the posterior PDF of the parameters A and B, assuming Φ_1 as the initial PF. The posterior PDF has a conspicuous maximizer, and the model-based moment ratio $\mathcal{K}(t)$ computed by using the maximizer successfully reproduced the real data-based moment ratio K(t), as shown in Fig. 10(b). Moreover, the posterior PDF shows clearly that a linear correlation between A and B exists. Recalling Eqs. (6) and (7), this linear structure implies that the uncertainty of W is small while that of γL is large. This is confirmed by the posterior PDFs of W and γL , which were obtained by applying a variable transformation to $p(A, B|D, \hat{\sigma}_1, \Phi_1)$ based on Eqs. (6) and (7):

$$p(W|\mathbf{D}, \hat{\sigma}_1, \Phi_1) = \int_{-\infty}^{\infty} dA \int_{-\infty}^{\infty} dB \delta \left(W - \pi \sqrt{\frac{2B}{A}} \right) \times p(A, B|\mathbf{D}, \hat{\sigma}_1, \Phi_1), \tag{23}$$

and

$$p(\gamma L|\mathbf{D}, \hat{\sigma}_1, \Phi_1) = \int_{-\infty}^{\infty} dA \int_{-\infty}^{\infty} dB \delta \left(\gamma L - \frac{\pi \sqrt{AB}}{4\sqrt{2}}\right) \times p(A, B|\mathbf{D}, \hat{\sigma}_1, \Phi_1). \tag{24}$$

Figures 10(c) and 10(d) show that the posterior PDF of W seems to be unimodal, while that of γL was multimodal. The relation between the properties of the MPF model and the data-based moment ratio K can explain the different features of these posterior PDFs. The MPF model [Eq. (5)] can be rewritten using W and γL :

$$\frac{\partial \phi_{i}}{\partial t} = 4\gamma L s_{i} \left[\underbrace{\frac{2}{W} \left(\phi_{i} - \frac{1}{n} \right)}_{\text{(i)}} + \underbrace{\frac{W}{\pi^{2}} \left(\nabla^{2} \phi_{i} - \frac{1}{n} \sum_{j=1}^{N} s_{j} \nabla^{2} \phi_{j} \right)}_{\text{(ii)}} \right]. \tag{25}$$

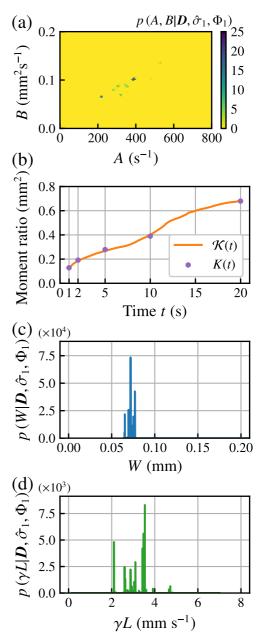


FIG. 10. Results of application to the real data-based moment ratio K(t). (a) Joint posterior PDF $p(A, B|D, \hat{\sigma}_1, \Phi_1)$. (b) Comparison between the data-based moment ratio K(t) (the same as in Fig. 3) and the model-based moment ratio $\mathcal{K}(t)$ computed by using the maximizer (\hat{A}, \hat{B}) of $p(A, B|D, \hat{\sigma}_1, \Phi_1)$. (c), (d) The marginal PDFs of W and γL transformed from $p(A, B|D, \hat{\sigma}_1, \Phi_1)$ via Eqs. (23) and (24).

In this equation, term (i) sharpens the grain boundaries, whereas term (ii) makes them diffuse. The magnitude relation between these terms determines the spatial functional form of the PFs in the steady state. When term (i) is too large, i.e., W is too small, the motion of all grain boundaries "freeze" before the curvature-driven migration of grain boundaries starts. When term (ii) is too large, i.e., W is too large, all grain boundaries diffuse; then the PFs cannot appropriately describe the grain structure. These facts suggest that an appropriate W exists for reproducing K, which is driven by grain-boundary motion.

On the other hand, the quantity γL is the characteristic rapidity of the change in the PFs. In the special case where the number of grains N is two, the migration speed of the grain boundary is proportional to γL . In this case, the determination of γL works well because the variation of \mathcal{K} is completely determined by the migration of the boundary. In fact, a previous study [27] showed that DA using the ensemble Kalman filter [14] could determine the PF mobility in a structural-phase-transition problem for steel. When N is greater than three, such as in the problem considered in this work, many simultaneous migrations of grain boundaries make the posterior PDF of γL broad and/or multimodal, as shown in Fig. 10(d). Nevertheless, obtaining more information to constrain the migration speed of grain boundaries remains an open problem.

When the proposed methodology based on the MPF model [Eq. (5)] is applied to materials that exhibit abnormal grain growth [4], unlike the steel used in this study, the posterior PDF variance will be larger than Fig. 10(a). To reduce this variance, a more accurate PF model that can appropriately reproduce the abnormal grain growth is required. Our methodology is available not only for the current MPF model [Eq. (5)] but also for other PF models involving more complex physical effects. Therefore, the best PF model can be selected through a comparison of the obtained posterior PDFs.

VII. CONCLUSIONS

We proposed a methodology based on Bayesian inference that enables us to estimate unobservable parameters involved in MPF models for grain growth. This enables accurate simulations that can properly explain real experimental data; such data-driven optimizations of massive simulations have previously proven difficult in structural materials science. The key points in our methodology are as follows: (i) the construction of the time-dependent data-based statistic from the static experimental data, which can be directly used in the procedure of DA, (ii) the EBM-based estimation of the posterior PDF without explicitly assuming a hyperparameter related to the observational noise, and (iii) the procedure of selecting the best initial PF that matches experimental data.

We validated the proposed methodology through three twin experiments using synthetic data and confirmed that the estimation of the parameters involved in the MPF model and the selection of the initial PF were successful. The proposed methodology was then applied to real experimental data related to the grain structures of a metal. The approach objectively estimated parameters along with their uncertainties by constructing the posterior PDF and selected the initial PF that best matched the time series of the data-based statistic. The obtained posterior PDF of the thickness of the grain boundary W seemed to be unimodal. This implied that an appropriate W for the MPF simulation existed, although W was a nonphysical quantity used only for PF simulations. On the other hand, the posterior PDF of the quantity γL related to the migration speed of the grain boundary seemed to be multimodal. This meant that more experimental data were needed to constrain the migration speed of the grain boundary.

This work focused on grain growth driven purely by the curvature of grain boundaries, which is the most fundamental effect in grain growth. However, the proposed EBM-based methodology is widely applicable to other PF models, including complex physical effects such as the anisotropy of boundary migration [5,34], thermal effects [9,34], elasticity [35,36], and transport of chemical compositions [6,7]. The proposed methodology can evaluate posterior PDFs related to unmeasurable or unobservable parameters originating from such effects. Such posterior PDFs provide us with valuable information not only for improving predictions in the dynamics of grain growth, but also for optimization of observational settings. This can be used to address problems such as how to assign time intervals for measurements, which is a crucial problem in practical experiments.

ACKNOWLEDGMENTS

This work was supported by the Council for Science, Technology and Innovation (CSTI), the Cross-ministerial Strategic Innovation Promotion Program (SIP) "Structural Materials for Innovation", funded by Japan Science and Technology Agency (JST). Development of part of the key algorithms was triggered by discussions in the research projects of JST CREST Grants No. JPMJCR1761 and No. JPMJCR1763, Japan Society for the Promotion of Science (JSPS) Grants-in-Aid for Scientific Research (B) Grants No. JP17H01703 and No. JP16H02823, Grant-in-Aid for Young Scientists Grants No. JP16K17779 and No. JP19K14671, and ERI JURP 2016-B-11 and 2018-B-01, The University of Tokyo, Japan.

P. A. Manohar, D. P. Dunne, T. Chandra, and C. R. Killmore, ISIJ Int. 36, 194 (1996).

^[2] S. Uhm, J. Moon, C. Lee, J. Yoon, and B. Lee, ISIJ Int. 44, 1230 (2004).

^[3] S.-J. Lee and Y.-K. Lee, Mater. Eng. (Reigate, UK) 29, 1840 (2008).

^[4] M. Hillert, Acta. Metall. 13, 227 (1965).

^[5] A. A. Wheeler, J. Stat. Phys. 95, 1245 (1999).

^[6] A. A. Wheeler, W. J. Boettinger, and G. B. McFadden, Phys. Rev. A 45, 7424 (1992).

^[7] D.-H. Yeon, P.-R. Cha, and J.-K. Yoon, Scr. Mater. 45, 661 (2001).

^[8] P. J. Alberry, B. Chew, and W. K. C. Jones, Met. Technol. (London) 4, 317 (1977).

^[9] M. Ohno, T. Yamaguchi, D. Sato, and K. Matsuura, Comput. Mater. Sci. 69, 7 (2013).

^[10] I. Steinbach and F. Pezzolla, Phys. D 134, 385 (1999).

^[11] I. Steinbach, F. Pezzolla, B. Nestler, M. Seeßelberg, R. Prieler, G. J. Schmitz, and J. L. L. Rezende, Phys. D 94, 135 (1996).

^[12] S. G. Kim, D. I. Kim, W. T. Kim, and Y. B. Park, Phys. Rev. E 74, 061605 (2006).

^[13] G. Welch and G. Bishop, An Introduction to the Kalman Filter, Tech. Rep. 95-041 (Dept. Computer Science, Chapel Hill, 1995).

- [14] G. Evensen, Ocean Dyn. 53, 343 (2003).
- [15] G. Kitagawa, Introduction to Time Series Modeling (Chapman & Hall/CRC, Boca Raton, 2010).
- [16] S. Reich and C. Cotter, *Probabilistic Forecasting and Bayesian Data Assimilation* (Cambridge University Press, Cambridge, 2015).
- [17] S. Ito, H. Nagao, A. Yamanaka, Y. Tsukada, T. Koyama, M. Kano, and J. Inoue, Phys. Rev. E 94, 043307 (2016).
- [18] G. E. P. Box and G. C. Tiao, *Bayesian Inference in Statistical Analysis*, Addison-Wesley Series in Behavioral Science: Quantitative Methods (Addison-Wesley, Reading, MA, 1973).
- [19] E. Kalnay, Atmospheric Modeling, Data Assimilation and Predictability (Cambridge University Press, Cambridge, 2003).
- [20] T. Tsuyuki and T. Miyoshi, J. Meteorol. Soc. Jpn. 85B, 331 (2007).
- [21] M. Ghil and P. Malanotte-Rizzoli, Adv. Geophys. 33, 141 (1991).
- [22] M. Kano, S. Miyazaki, Y. Ishikawa, Y. Hiyoshi, K. Ito, and K. Hirahara, Geophys. J. Int. 203, 646 (2015).
- [23] R. Niwayama, H. Nagao, T. S. Kitajima, L. Hufnagel, K. Shinohara, T. Higuchi, T. Ishikawa, and A. Kimura, PLoS One 11, e0159917 (2016).
- [24] T. Kuwatani, H. Nagao, S. Ito, A. Okamoto, K. Yoshida, and T. Okudaira, Phys. Rev. E 98, 043311 (2018).

- [25] T. Koyama, Y. Tsukada, and Y. Kawai, in *Proceedings of the International Conference on Solid-Solid Phase Transformations in Inorganic Materials* 2015 (PTM 2015, Whistler, British Columbia, 2015), pp. 831–832.
- [26] S. Ito, H. Nagao, T. Kasuya, and J. Inoue, Sci. Tech. Adv. Mater. 18, 857 (2017).
- [27] K. Sasaki, A. Yamanaka, S. Ito, and H. Nagao, Comput. Mater. Sci. 141, 141 (2018).
- [28] D. L. Olmsted, J. L. G. Hector, W. A. Curtin, and R. J. Clifton, Modell. Simul. Mater. Sci. Eng. 13, 371 (2005).
- [29] D. L. Olmsted, E. A. Holm, and S. M. Foiles, Acta Mater. 57, 3704 (2009).
- [30] R. D. Kamachali, A. Abbondandolo, K. F. Siburg, and I. Steinbach, Acta Mater. 90, 252 (2015).
- [31] G. Casella, Am. Stat. 39, 83 (1985).
- [32] B. P. Carlin and T. A. Louis, *Bayes and Empirical Bayes Methods for Data Analysis*, 2nd ed., Texts in Statistical Science (Chapman & Hall/CRC, Boca Raton, 2000).
- [33] H. Nagao and T. Higuchi, in *Proceedings of the Institute of Statistical Mathematics* (The Institute of Statistical Mathematics, Japan, 2013), Vol. 61, pp. 257–270.
- [34] R. Kobayashi, Phys. D 63, 410 (1993).
- [35] L.-Q. Chen, Annu. Rev. Mater. Res. 32, 113 (2002).
- [36] M. Tonks and P. Millett, Mater. Sci. Eng., A 528, 4086 (2011).