

Full length article

Uncertainty propagation in a multiscale CALPHAD-reinforced elastochemical phase-field model

Vahid Attari^{a,*}, Pejman Honarmandi^a, Thien Duong^a, Daniel J. Saucedo^a, Douglas Allaire^a, Raymundo Arroyave^{a,b}

^a Materials Science and Engineering Department, Texas A&M University, College Station, TX 77840, USA

^b Mechanical Engineering Department, Texas A&M University, College Station, TX 77840, USA

ARTICLE INFO

Article history:

Received 9 July 2019

Revised 9 November 2019

Accepted 11 November 2019

Available online 18 November 2019

Keywords:

Phase-field modeling

Uncertainty propagation

Uncertainty quantification

Thermoelectrics

Microstructure

Mass scattering

Phonon scattering

ABSTRACT

ICME approaches provide decision support for materials design by establishing quantitative process-structure-property relations. Confidence in the decision support, however, must be achieved by establishing uncertainty bounds in ICME model chains. The quantification and propagation of uncertainty in computational materials science, however, remains a rather unexplored aspect of computational materials science approaches. Moreover, traditional uncertainty propagation frameworks tend to be limited in cases with computationally expensive simulations. A rather common and important model chain is that of CALPHAD-based thermodynamic models of phase stability coupled to phase-field models for microstructure evolution. Propagation of uncertainty in these cases is challenging not only due to the sheer computational cost of the simulations but also because of the high dimensionality of the input space. In this work, we present a framework for the quantification and propagation of uncertainty in a CALPHAD-based elastochemical phase-field model. We motivate our work by investigating the microstructure evolution in $\text{Mg}_2\text{Si}_x\text{Sn}_{1-x}$ thermoelectric materials. We first carry out a Markov Chain Monte Carlo-based inference of the CALPHAD model parameters for this pseudobinary system and then use advanced sampling schemes to propagate uncertainties across a high-dimensional simulation input space. Through high-throughput phase-field simulations we generate 200,000 time series of synthetic microstructures and use machine learning approaches to understand the effects of propagated uncertainties on the microstructure landscape of the system under study. The microstructure dataset has been curated in the Open Phase-field Microstructure Database (OPMD), available at <http://microstructures.net>.

© 2019 Acta Materialia Inc. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Uncertainty Quantification (UQ) has a long and successful history of application to very diverse areas such as climate change [1], structural engineering [2], aerospace engineering and design [3], and medicine [4], to name a few. In the field of materials science, however, notions of UQ remain relatively unexplored even though proper quantification of uncertainty in models and simulations is of critical importance as the field progresses towards more quantitative/predictive approaches to materials discovery and development. Indeed, uncertainty quantification (UQ) and its propagation (UP) across model/simulation chains are considered key elements of decision-based [5–8] materials design in the framework of Integrated Computational Materials Engineering (ICME) [9]. The latter

prescribes the integration of databases, multi-scale modeling and experiments with the aim of reducing the time and effort of the materials development cycle [10]. Given the complexity and computational cost of most materials simulation frameworks, it is necessary to have a systematic and efficient approach to quantify uncertainties in the parameters/variables in any system of interest and to propagate these uncertainties to the respective responses of individual or multi-scale systems.

Despite the importance of UQ/UP in multi-scale modeling [11,12], there are very few works in the literature dealing with UQ and/or UP across multi-scale models in the field of materials science and engineering. Liu et al. [13], for example, focused on the probabilistic prediction of the effective properties in heterogeneous composite materials and their performance. In that work, UQ of the parameters and UP across the multi-scale constitutive models (i.e. UP from structure to property to performance) were performed through a Bayesian stochastic method and a stochastic projection technique, respectively. Some works for UP across multiple

* Corresponding author.

E-mail address: attari.v@tamu.edu (V. Attari).

scales for the probabilistic predictions of plastic flow behavior in poly-crystalline materials have been described thoroughly in [14–16]. Recently, Honarmandi et al. [17] demonstrated the use of UQ approaches to the parameterization of thermodynamically rigorous models for the response of NiTi-based shape memory alloys, followed by uncertainty propagation over the model parameter space.

In computational materials science, thermodynamic assessments using the CALculation of PHase Diagrams (CALPHAD) method [18] constitutes the basis for a broad range of approaches to materials simulations, including microstructure evolution through phase-field modeling [19]. Given the foundational nature of CALPHAD-based descriptions of phases' free energies in any attempt to predict processing-(micro)structure relationships, UQ/UP in CALPHAD [20] play a very important role, although sparse examples in the literature address this. Honarmandi et al. [21] used a Bayesian framework to quantify and propagate uncertainty in the context of CALPHAD thermodynamic assessments and showed how information fusion approaches [22] can be used to fuse propagated uncertainties from different competing models. Other groups have also demonstrated different frameworks for the quantification and propagation of uncertainty in CALPHAD models [23–25].

Similarly to the case of CALPHAD-based thermodynamic assessments, the application of UQ/UP frameworks to phase-field modeling remains relatively unexplored [26–31]. Koslowski et al. [15] characterized how uncertainties propagate across spatial and temporal scales in a physics-based model of nanocrystalline plasticity of fcc metals, combining molecular dynamics (MD) with phase-field dislocation dynamics (PFDD) simulations. Wang et al. [32] carried out an asymptotic uncertainty analysis of void formation of materials under irradiation conditions. Leon et al. [33] used subset selection and active subspace techniques to identify dominant parameters in a continuum phase-field polycrystalline model for ferroelectric materials. While these earlier approaches focused on UQ/UP over a single modeling framework, Böttger [34] recently demonstrated the propagation of uncertainty across an entire ICME-based model chain.

Across different fields [35], UP is practically implemented through different approaches, including Monte Carlo (MC)-, local expansion-, functional expansion-, and numerical integration-based methods. By far, the most basic and common approach to propagating uncertainty through computational models is via MC simulation [36]. For expensive computational models, however, the use of MCMC sampling-based approaches is often computationally prohibitive. While one can use surrogate models to efficiently sample the input/output relationships in simulations [37], numerical efficiency often comes at the expense of fidelity—i.e. the generation of surrogate models more often than not involves a considerable reduction in the dimensions of the output space. Moreover, such approaches tend to fail in cases in which the model output changes qualitatively (not only quantitatively) in different regions of the input space—i.e. when the physical response of the system under study undergoes abrupt transitions over relatively short distances in the input space. Phase-field simulations belong to the latter class of computational problems where these conventional approaches to UP tend to be ineffective—e.g. in phase-field models solidification small changes in input parameters and initial conditions can lead to transitions from columnar to equiaxed growth [38]. The challenges associated with the computational expense and the non-regular nature of the output of phase-field simulations are augmented by the large dimensionality of the input space. Finally, the complexity of the output of phase-field simulations makes the analysis of their propagated uncertainties quite challenging.

The major objective of the current work is to show the propagation of statistically quantified uncertainties of thermodynamic

parameters to the Gibbs free energies of phases, and equilibrium phase diagram, in combination with the uncertainties of microelastic and kinetic parameters of an elastochemical phase field model. Consequently, uncertainties in the Gibbs free energies and their propagation to microstructural characteristics are studied by using high throughput analyses across a chain of simulation tools that includes a CALPHAD, a microelasticity, and a phase-field model, consecutively. The uncertainty in thermodynamic parameters in the CALPHAD model is quantified through Markov Chain Monte Carlo (MCMC) sampling in the context of Bayesian statistics, while the uncertainty of microelastic and kinetic parameters is determined through prior knowledge, calculations and/or expert judgement-based estimations.

Fig. 1 demonstrates the proposed strategy and steps toward developing a framework for propagating the uncertainty across CALPHAD reinforced phase-field model chains. The steps are based on determining 1) type of processing conditions, 2) interacting physics (sub-models), 3) Quantities of Interest (QoIs), and 4) microscopic/macroscale properties of interest. We demonstrate the framework through the investigation of the microstructure evolution of nanostructured $\text{Mg}_2\text{Si}_x\text{Sn}_{1-x}$ thermoelectric (TE) materials [39,40]. This pseudo-binary system is characterized by a miscibility gap between two isomorphous cubic phases [40]. The tendency of this system to phase-separate or homogenize is taken into account by employing a fully parameterized elastochemical phase-field model that accounts for the effect of process conditions on the resulting microstructure configurations.

Since the performance of multi-phase thermoelectric materials is greatly determined by their (multi-scale) microstructure, it is expected that changes in processing schemes may have considerable impact on performance. In fact, for the case of $\text{Mg}_2\text{Si}_x\text{Sn}_{1-x}$ -based thermoelectrics, it has been already shown that equilibrium and non-equilibrium processing can lead to dramatic changes to the TE figure of merit, zT [40–42]. Understanding of this TE system is further hindered by the considerable uncertainty in the location of the miscibility gap, with different experimental phase boundary estimations disagreeing by several tens of atomic percent [40].

In this work, we attempt to propagate uncertainty through phase-field simulations, accounting for the uncertainty not only in the parameters directly associated with the phase-field model, but also arising from uncertainty in the CALPHAD parameterization. This paper is structured as follows: in Section 2, we motivate the present work through the application of ICME-based frameworks on the design of nanostructured TE materials. We note that the framework put forward is generalizable to a wide range of materials simulation problems. The details of the models (CALPHAD and phase-field) and uncertainty propagation strategy is provided in Section 3. Section 4 discussed the prior and the resulting uncertainty in the phase diagram, the subsequent uncertainty in the microstructures under elastochemical phase field simulations, and the methods for data interpretation/classification. Furthermore, we present a summary on our findings and draw our conclusions in Section 5.

2. Nanostructured thermoelectric materials

As mentioned above, the motivating example in this work is the propagation of uncertainty in microstructure evolution simulations of nanostructured composite thermoelectrics (TE) [43–46] via direct coupling of CALPHAD thermodynamic databases [47] with multi-physics phase-field models (PFM) [28,40] that account for both chemical and elastic driving forces for structure formation. The example is motivated by recent work by some of the present authors [40] on the dramatic effect that processing has on the microstructure (and TE performance) in $\text{Mg}_2\text{Si}_x\text{Sn}_{1-x}$ alloys, but has a much broader applicability as CALPHAD/PFM-based mi-

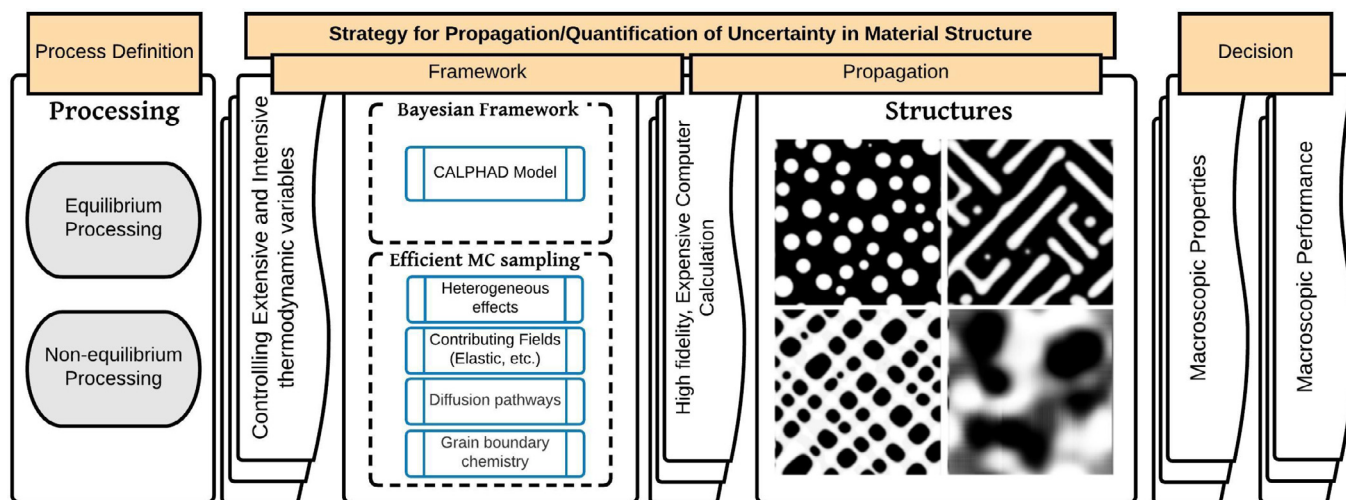


Fig. 1. The process-structure-property uncertainty propagation framework deployed in composition, strain and temperature space based on the process definition and natural uncertainties in input parameters. Updated Figure.

crostructure simulations are pervasive in ICME-based frameworks for microstructure-sensitive materials design [48–52], and properly accounting for uncertainty along processing-(micro)structure-properties-performance relationships is necessary to make materials design choices with proper confidence bounds.

Current interest in the thermoelectric (TE) effect originates from the ever increasing demand for energy and the associated detrimental effects on global climate. Current TE materials, unfortunately, do not have the efficiency—quantified by the figure of merit $zT = \frac{\sigma_e S^2 T}{\kappa}$, where σ_e is the electrical conductivity, S is the Seebeck coefficient, T is the absolute temperature and κ is the thermal conductivity—that would turn TE-based devices into competitive power-generators [43,53]. An ideal TE material would have a large Seebeck coefficient, while being electrically conductive and thermally insulating [43,53]. These properties, however, are strongly coupled and their individual tuning is thus challenging. Over the past decade, a sophisticated arsenal of strategies for the rational design of TE materials has emerged [54–56], including the exploitation of spontaneous self-assembly or non-equilibrium processing of nanostructures to enhance phonon-scattering [40,43].

Among the hundreds of TE systems investigated to date, environmentally-benign $\text{Mg}_2\text{Si}_x\text{Sn}_{1-x}$ alloys [57,58] have attracted considerable attention due to their relatively high figure of merit ($zT > 1$) [59], comparable with intermediate temperature TE materials such as PbTe and filled skutterudites [40,58,60,61]. The $\text{Mg}_2\text{Si}_x\text{Sn}_{1-x}$ pseudo-binary system exhibits a miscibility gap [62–65] and this has been exploited to realize nanostructures with optimal TE performance [57,59,66]. Experimental determination of zT even in a single alloy exhibits considerable scatter, perhaps due to (subtle) changes in the way these materials are synthesized and processed. Recently, the present authors and collaborators investigated the effect of non-equilibrium processing on the microstructure evolution (and transport properties) in the $\text{Mg}_2\text{Si}_{0.7}\text{Sn}_{0.3}$ system and found that instead of phase-separating, the system tended to form a solid-solution with superior TE performance, contrary to expectations and prior works [57,66]. This was ascribed to (elastic) coherency effects and was verified via quantitative multi-physics phase-field simulations [40].

These results are interesting as they exemplify the influence of processing on the microstructural evolution in TE materials [43] and the corresponding change in performance. Most importantly, this constitutes one of the very few examples—to the best of our knowledge—in which phase-field modeling (PFM) in

combination with CALPHAD free energies has been used to investigate the microstructure evolution of TE materials. Further investigation of the PFM developed to study the $\text{Mg}_2\text{Si}_{0.7}\text{Sn}_{0.3}$ showed that rather small changes in the strength of the elastic couplings—mediated via lattice parameter differences between Si- and Sn-rich domains—resulted in qualitatively different microstructures, which in turn could be expected to exhibit different phonon transport behavior.

The ability to quantitatively understand [43] and control the different materials and processing parameters related to microstructural morphology, topology, size and spacing in composite TE materials has already been demonstrated [46,67–70]. Many of these approaches have been inspired by metallurgy and thus the time is ripe to translate much of what has been learned on ICME-enabling microstructure-sensitive (structural) alloy design to the problem of designing (self-assembled) TE microstructures for optimal performance. While the modeling framework via PFM has shown to result in (semi-)quantitative predictions that compare well with experiments [40], a robust ICME research program on microstructure design of TE materials requires reliable and efficient UQ/UP frameworks.

Fig. 2 illustrates an schematic phase diagram in which the material shows an inherent chemical instability in certain regions of the composition space. This material is uniform and homogeneous at high temperatures and upon reducing the temperature spontaneously decomposes into distinct phases. Similarly, the $\text{Mg}_2\text{Si}_x\text{Sn}_{1-x}$ system has a miscibility where the uncertainties in the Sn-rich and Si-rich boundaries (refer to Fig. 3 of [40]) has been the center of discussions in several prior works [39,40,62–64]. The uncertainties on the location of these phase boundaries necessarily impact the chemical free energy and (the predicted) microstructure of this system.

The sensitivity of the microstructural evolution of these systems to chemical instabilities, the uncertainty in the boundaries of the unstable region and the wide range of synthesis/processing conditions all suggest that analysis and quantification of uncertainty in quantitative microstructure simulation models is necessary in this material to establish the correlation between the thermodynamics of the system and the resulting microstructural phenomena. These analyses are essential to estimate how the variance in processing and (epistemic or aleatoric) uncertainties in material parameters can affect the microstructure and ultimately the TE response of this material. This also can contribute to a better identification of

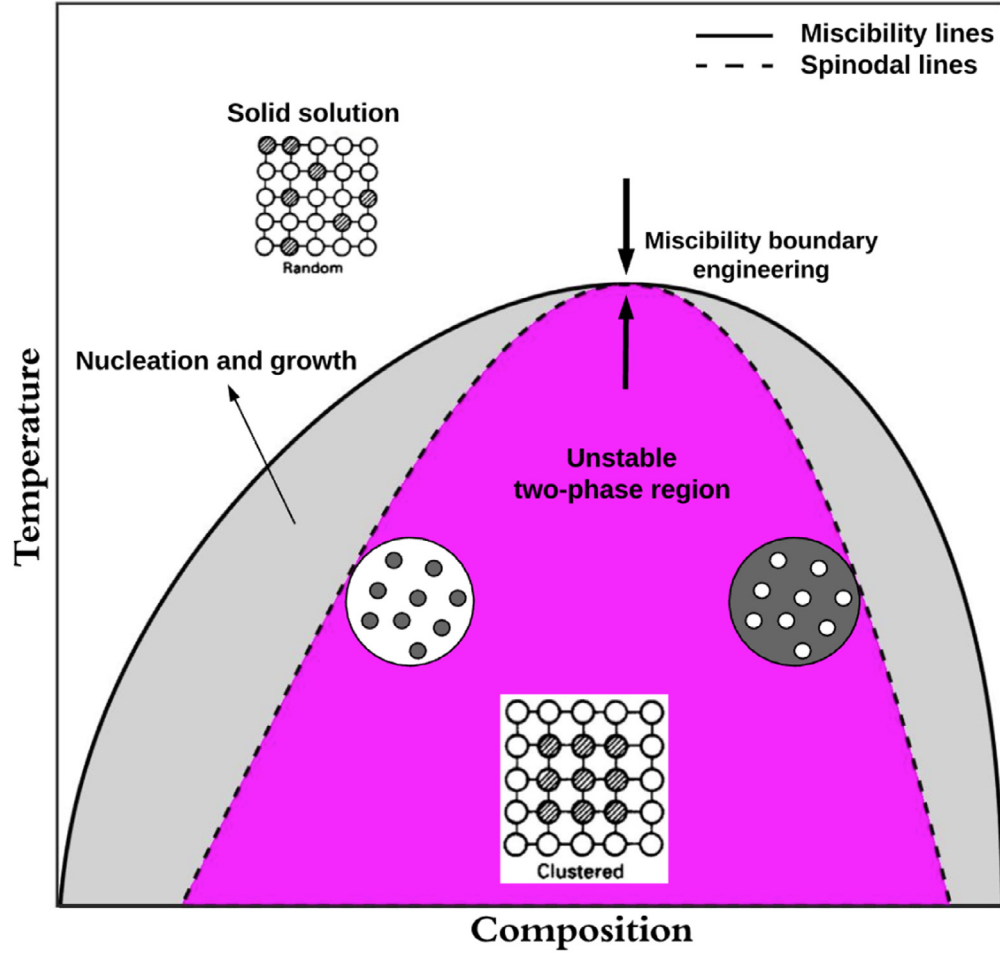


Fig. 2. Schematic phase diagram for an imaginary binary alloy exhibiting a chemical spinodal. A strategy to enhance thermoelectric performance is engineering the spinodal line location by non-equilibrium processing for better phonon scattering via mass scattering versus nanoscale precipitates which in turn cause interface scattering. Updated figure.

the optimal thermodynamic and processing conditions for the enhancement of the properties of TE materials [71].

3. Models and methodologies

Here, we first define the thermodynamic state variables relevant to the elastochemical phase-field model formulation and then quantify the uncertainties in the given material variables and propagate these uncertainties with the aim of quantifying the variations in the micro-structure of the material that ultimately alters the predicted macroscopic response.

3.1. Thermodynamic state variables

For an isothermal and isobaric state, the total free energy functional ($\mathcal{F}^{tot}$) for an undeformed material configuration with confined boundaries can be constructed from the sum of all contributing fields over it. We restrict $\mathcal{F}^{tot}$ to chemical and elastic contributing effects here. Consequently, the free energy of the microstructure is a functional of composition (c), temperature (T), strain (ε), gradients of composition (∇c), and other fields, if they were present. We write $\mathcal{F}^{tot}$ as the sum of all contributing fields for a confined volume (Ω) as:

$$\mathcal{F}^{tot}[c, T, \nabla c, \varepsilon^{el}] = \int_{\Omega} [f_{bulk} + f_{interfacial} + f_{elastic}] d\Omega \quad (1)$$

where bulk free energy, f_{bulk} , interfacial free energy, $f_{interfacial}$, and elastic strain energy, $f_{elastic}$ are:

$$f_{bulk} = f^0(c, T) \quad (2)$$

$$f_{interfacial} = \kappa (\nabla c)^2 \quad (3)$$

$$f_{elastic} = \frac{1}{2} \sigma_{ij} \varepsilon_{ij}^{el} \quad (4)$$

where $f^0(c, T)$ is the free energy of a unit volume of homogeneous material, κ is the gradient energy coefficient, ε_{ij}^{el} and σ_{ij} are the local elastic strain and stress in the material, respectively. The chemical free energy is composed of the interfacial and bulk energy contributions, and it determines the compositions and volume fractions of the equilibrium phases. The strain energy affects the equilibrium compositions and volume fractions of the coexisting phases, as well as the shapes and configurations of the phase domains.

The bulk free energy of the $\text{Mg}_2\text{Si}_x\text{Sn}_{1-x}$ pseudo-binary system is described through the sub-regular solution model as:

$$f_{bulk}^{\phi}(c, T) = \sum_{\alpha} c_{\alpha} \cdot {}^0 G_{\alpha}^{\phi} + RT \sum_{\alpha} c_{\alpha} \ln(c_{\alpha}) + \sum_{\alpha} \sum_{\beta \neq \alpha} c_{\alpha} c_{\beta} \sum_{\nu} {}^{\nu} L_{\alpha, \beta}^{\phi} (c_{\alpha} - c_{\beta})^{\nu} \quad (5)$$

where α (β) represents the constituents of the alloy, and ϕ denotes Mg_2Si , Mg_2Sn , or liquid phases. $^0G_\alpha^\phi$ is the reference energy of the constituent, T is temperature (in K), R is the universal gas constant, and $^vL_{\alpha\beta}^\phi$ is Redlich–Kister coefficient of the expansion order v given by:

$$^vL_{\alpha,\beta}^\phi = ^va_{\alpha,\beta}^\phi + ^vb_{\alpha,\beta}^\phi T \quad (6)$$

where $^va_{\alpha,\beta}^\phi$ and $^vb_{\alpha,\beta}^\phi$ are model parameters that describe the interactions between the constituents beyond those corresponding to ideal mixtures. These parameters can be calibrated against the available data deterministically or probabilistically, as will be discussed later. In order to build the phase diagram, the total Gibbs free energy of the system is minimized at different temperatures for volume fraction and composition of existing phases. The two phases Mg_2Si and Mg_2Sn have the same crystal structure (CaF₂ structure type, cF12 Pearson symbol).

3.2. Material modeling strategy: CALPHAD reinforced phase-field method

From linear kinetic theory, the local mass flux in the presence of a gradient in composition, the diffusion flux, $\mathbf{J}$ (in units of $\text{mol m}^{-2} \text{s}^{-1}$) is given by:

$$\mathbf{J} = -\mathbf{M}\nabla\mu^{\text{tot}} \quad (7)$$

where $\mathbf{M}$ is the interface mobility taken to be equal in all directions due to the isotropic nature of the crystal structures of the two phases, and $\mu^{\text{tot}} = \frac{\delta\mathcal{F}^{\text{tot}}}{\delta c}$ is the total potential for the kinetic transition, where δ is the variational derivative operator. We postulate the following form of the Cahn–Hilliard (C-H) kinetic equation along with other micro-elasticity equations to study the evolution of the $\text{Mg}_2\text{Si}_x\text{Sn}_{1-x}$ microstructure:

$$\frac{\partial c}{\partial t} = \nabla \cdot \mathbf{M}\nabla \left(\frac{\delta\mathcal{F}^{\text{tot}}}{\delta c} \right) \quad (8)$$

This equation is solved by utilizing a semi-implicit Fourier spectral approach [72] in frequency (k) space. The algorithm is provided in Appendix A. The simulation cell size is 512×512 in the entire study. The composition profile is perturbed $\pm 2\%$ randomly around the alloy composition with a constant seed number for all calculations to maintain consistency. We use a constant time step for the entire calculations, and the results are saved with a constant interval to make them comparable. This is important due to the fact that a change in kinetic parameters affects the coarsening rate, and stability of the simulations extensively. The micro-elasticity equations are given as:

$$\frac{\partial \sigma_{ij}}{\partial r_j} = 0 \quad \text{in } \Omega \quad (9)$$

$$\varepsilon_{ij} = \frac{1}{2} \left(\frac{\partial u_i}{\partial r_j} + \frac{\partial u_j}{\partial r_i} \right) \quad (10)$$

$$\sigma_{ij} = C_{ijkl} \varepsilon_{kl}^{\text{el}} \quad (11)$$

$$\varepsilon_{kl}^{\text{el}} = \varepsilon_{kl}^{\text{tot}} - \varepsilon_{kl}^0 \quad (12)$$

where Eqs. (9)–(12) are the mechanical equilibrium condition, kinematics, Hooke's microscopic constitutive law for linear elasticity, and strain relationship, respectively. u is the displacement field and r is the space vector. The dilatational eigenstrain term is given by $\varepsilon_{kl}^0 = \varepsilon^T \delta_{kl} h(c)$, and it is the consequence of lattice strain between the phases. ε^T is the strength of the eigenstrain, δ_{kl} is the Kronecker-delta function, and $h(c)$ is an standard scalar-valued interpolation function defined by $h(c) = c^3(10 - 15c + 6c^2)$. C_{ijkl} is

the composition-dependent fourth order elastic modulus tensor. It is convenient to describe C_{ijkl} using the following expression:

$$C_{ijkl}(c) = C_{ijkl}^{\text{eff}} - g(c) \Delta C_{ijkl} \quad (13)$$

where $\Delta C_{ijkl} = C_{ijkl}^\alpha - C_{ijkl}^\beta$ is the difference between the elastic moduli tensor of the α and β phases. For a linear elastic material, the general form of the C_{ijkl} matrix will be reduced to:

$$C_{ijkl} = \begin{bmatrix} C_{1111} & C_{1122} & C_{1133} & C_{1123} & C_{1113} & C_{1112} \\ C_{2211} & C_{2222} & C_{2233} & C_{2223} & C_{2213} & C_{2212} \\ C_{3311} & C_{3322} & C_{3333} & C_{3323} & C_{3313} & C_{3312} \\ C_{2311} & C_{2322} & C_{2333} & C_{2323} & C_{2313} & C_{2312} \\ C_{1311} & C_{1322} & C_{1333} & C_{1323} & C_{1313} & C_{1312} \\ C_{1211} & C_{1222} & C_{1233} & C_{1223} & C_{1213} & C_{1212} \end{bmatrix} \quad (14)$$

In a real case, the 21 independent constants of this matrix are simplified to 2, 5, or 9 depending on the material symmetry. In this study, we consider only 2 independent constants due to the cubic symmetry of the phases. Similarly to $h(c)$, $g(c) = h(c) - \frac{1}{2}$ is also an interpolation function that sets the average of the elastic constants of the two phases as the effective elastic tensor. The microelasticity problem (Eqs. 9–12) is solved by the fast Fourier transform (FFT)-based iterative method. For a detailed description of this approach, the reader is invited to refer to our previous work on microelasticity effects on microstructural evolution of $\text{Ti}_{1-x-y}\text{Al}_x\text{Zr}_y\text{N}$ alloy [29]. In summary, the following conditions are considered for the microelasticity model: 1) Stress-free transformation strains (SFTS) for each phase and inhomogeneous elastic constants in the domain, 2) $\Delta C_{ijkl} = 0$ in the first iteration, 3) strain-control based on stress-control, 4) convergence of the problem when the L^2 norm of $\|u^{n+1} - u^n\|$ is less than 10^{-8} , and 5) Periodic boundary conditions on all sides of the domain.

The parameter space for CALPHAD, micro-elasticity, and phase-field models is listed in Table 1.

3.3. Uncertainty quantification/propagation

The quantification of uncertainties associated with model parameters and predictions is one of the most important tasks in simulation-aided materials design. In this regard, quantifying the uncertainties of the model parameters given the available data is an inverse problem, known as (inverse) UQ, while propagating these uncertainties across an individual model or a chain of models is a forward problem known as (forward) UP.

Generally, UQ of model parameters can be performed in the context of either frequentist or Bayesian statistical inference. In this work, the latter has been applied to probabilistically calibrate the relevant model parameters. Here, the parameter calibration or uncertainty quantification is performed based on Bayes' rule where the parameter posterior probability given the available data is proportional to the parameter prior probability times likelihood, i.e. $P(\theta|D) \propto P(\theta)P(D|\theta)$ assuming θ and D are the parameters and data, respectively. Here, the prior probability distributions of the model parameters—which is defined through the available prior knowledge about the parameters—is updated to their posterior probability distributions given the data using the likelihood distribution—which is a conditional probability of the data given the model parameters. Bayesian inference thus results from the combination of prior beliefs and likelihood, as opposed to frequentist approaches in which the inference is just based on the likelihood.

The resulting posterior distribution represents the plausible optimal value of the parameters and their uncertainties that can be propagated to the model predictions through different analytical or numerical approaches. To find the statistical properties (mean

Table 1

First-order statistics and Probability Distribution Functions (PDFs) of CALPHAD model parameters and phase field model inputs. The outputs of the CALPHAD model were fed to the phase-field model. The fitted truncated PDFs correspond to the plausible optimal values and uncertainty of the model parameters. “ss” and “liq” denote $\text{Mg}_2\text{Si}_x\text{Sn}_{1-x}$ solid solution and liquid phases, respectively. ** SFTS stands for Stress-Free Transformation Strain.

Model	Sub-model	Parameter	PDF form	Unit	μ	Dispersion ($\frac{\sigma^2}{\mu}$)	Lower Bound	Upper Bound
Phase-field Model Parameters	CALPHAD Model Parameters	$^0G_{\text{Mg}_2\text{Sn},\text{Mg}_2\text{Si}}^{\text{ss}}$		J.mol^{-1}	12840.4	729.0	6824.89	20474.69
		$^0b_{\text{Mg}_2\text{Sn},\text{Mg}_2\text{Si}}^{\text{ss}}$		$\text{J.mol}^{-1}\text{K}^{-1}$	7.20	0.59	3.67	11.02
		$^1G_{\text{Mg}_2\text{Sn},\text{Mg}_2\text{Si}}^{\text{ss}}$		J.mol^{-1}	-3324.3	-288.3	-5208.34	-1736.11
		$^0G_{\text{Mg}_2\text{Sn},\text{Mg}_2\text{Si}}^{\text{liq}}$		J.mol^{-1}	80635.7	26688.6	43550.19	130650.57
	Phase-field Model Parameters	$^1b_{\text{Mg}_2\text{Sn},\text{Mg}_2\text{Si}}^{\text{liq}}$		$\text{J.mol}^{-1}\text{K}^{-1}$	-61.25	3.98	-86.03	-28.68
		$^1G_{\text{Mg}_2\text{Sn},\text{Mg}_2\text{Si}}^{\text{liq}}$		J.mol^{-1}	6400.5	529.01	3314.80	9944.41
	Kinetic Parameters	Alloy composition (c)		mol	0.40	0.0083	0.3	0.5
		Interface mobility (M)		$\text{m}^2\text{s}^{-1}\text{J}^{-1}$	5.62×10^{-19}	4.64×10^{-25}	$10^{-18}/(RT)$	$10^{-20}/(RT)$
		Gradient energy coefficient (κ)		J.m^{-2}	1.16×10^{-24}	4.64×10^{-25}	2.0×10^{-26}	2.0×10^{-24}
		SFTS ** (ϵ^T)		-	-3.1×10^{-5}	-4.28	-0.02	+0.02
	Microelasticity Model Parameters	$C_{11} \text{ Mg}_2\text{Sn}$		GPa	76.56	0.20	68.30	83.71
		$C_{12} \text{ Mg}_2\text{Sn}$		GPa	27.75	1.04	17.68	39.79
		$C_{44} \text{ Mg}_2\text{Sn}$		GPa	29.94	1.26	16.03	41.94
		$C_{11} \text{ Mg}_2\text{Si}$		GPa	120.15	0.07	114.07	126.00
		$C_{12} \text{ Mg}_2\text{Si}$		GPa	22.62	0.10	19.56	26.00
		$C_{44} \text{ Mg}_2\text{Si}$		GPa	46.79	0.63	33.32	58.20
		Molar volume ($V_m^{\text{Mg}_2\text{Si}}$)		$\text{m}^3.\text{mol}^{-1}$	5.78×10^{-5}	2.40×10^{-7}	4.73×10^{-5}	6.38×10^{-5}
		Molar volume ($V_m^{\text{Mg}_2\text{Sn}}$)		$\text{m}^3.\text{mol}^{-1}$	4.88×10^{-5}	1.55×10^{-7}	3.95×10^{-5}	5.33×10^{-5}

value and covariance matrix) of the posterior probability distribution, some intractable integrals should be solved. For this reason, MC-based approaches are often used to estimate these integrals [73,74]. In order to obtain well-converged estimates for the relevant integrals, a substantial number of evaluations of the model must be made. Generally, MCMC approaches are the most commonly used sampling techniques in the context of Bayesian inference. Among these techniques, the Metropolis-Hastings (MH) algorithm has been selected in this work, as described in detail in Section 3.3.1.

We note that it is also important to propagate uncertainties from the parameters to the model predictions since these probabilistic predictions of the given system are the quantities of interest in materials design. Three commonly used approaches for UP are analytical methods, numerical MC methods, and stochastic finite element analysis. In analytical methods, the output/prediction uncertainties are represented explicitly as functions of parameter uncertainties. These methods are useful when only small ranges of uncertainty are considered and cannot always be applied to

complex, nonlinear models with large parameter uncertainties. MC methods involve the estimation of PDFs for selected model outputs by performing a sufficiently large number of model runs with randomly sampled parameters. For computationally intensive models, the time and resources required for MC methods could be prohibitively expensive. A degree of computational efficiency is accomplished by the use of Modified Monte Carlo (MMC) methods that sample from the parameter distribution in an efficient manner so that the number of necessary samples are significantly reduced compared to simple MC methods. Such methods include the Fourier Amplitude Sensitivity Test (FAST) [75] and Latin Hypercube Sampling [76]. However, even these methods require a substantial number of model runs in order to obtain a good approximation of the output PDFs, especially for cases involving a large number of uncertain parameters. Therefore, there is a need for even more computationally efficient ways for propagating uncertainty in complex/expensive models and this is currently being investigated by the present authors.

In this work, a forward UP analysis based on optimal sampling methods has been applied to propagate the uncertainties across the chain of models, which is explained in Section 3.3.2.

3.3.1. Applied MCMC approach for the parameter uncertainty quantification

The MCMC toolbox in Matlab [77] has been utilized to calibrate the Gibbs free energy parameters of a CALPHAD model for the pseudo-binary $\text{Mg}_2\text{Si}_x\text{Sn}_{1-x}$ system. Since there is no prior knowledge about the distribution of the parameters and their correlations, a uniform prior probability distribution has been considered for the model parameters in this work. However, the initial values (θ^0) and ranges of the parameters have been determined based on the deterministically calibrated parameters obtained using the Thermo-Calc software in order to make the parameter convergence faster during MCMC sampling. After defining the prior probability distribution, the values of the parameters are randomly sampled from an arbitrary posterior proposal distribution (q) iteratively. In this work, the proposal distribution is an adaptive multi-variate Gaussian distribution that is centered at the last accepted parameter vector in the MCMC chain with an adaptive covariance matrix during MCMC sampling based on the resulting covariance from the previous parameters in the chain, according to Haario et al. [78]. In each iteration, the sampled candidate is accepted or rejected based on the Metropolis-Hastings ratio:

$$M - H = \frac{P(\theta^{\text{cand}})P(D|\theta^{\text{cand}})}{P(\theta^{z-1})P(D|\theta^{z-1})} \frac{q(\theta^{z-1}|\theta^{\text{cand}})}{q(\theta^{\text{cand}}|\theta^{z-1})} \quad (15)$$

This ratio compares the posterior probability of the sampled candidate (θ^{cand}) with its counterpart for the last accepted parameter vector in the MCMC chain (θ^{z-1}) through the metropolis ratio (the first ratio in Eq. (15)), and also compares the probability of moving from θ^{z-1} to θ^{cand} with the probability of the reverse move through the Hastings ratio (the second ratio in Eq. (15)) in the case that the posterior proposal distribution is asymmetric. It should be noted that the likelihood function is also a Gaussian distribution centered at the given data (D) with a variance that corresponds to the uncertainty of the data. Since the uncertainty of the data used in this work is not properly understood, an unknown variance has been considered for the likelihood which has been updated as a hyper-parameter with the model parameters during MCMC sampling—this has been explained in more detail by Gelman et al. [73]. $\min(M - H, 1)$ indicates the acceptance probability of the candidate in each iteration. θ^z equals θ^{cand} in the case of acceptance which updates the mean value of the proposal distribution; while it equals θ^{z-1} if the candidate is rejected. The sampling process proceeds until convergence of the MCMC chain to a stationary distribution, which represents the parameter convergence during the MCMC process. After discarding the burn-in period which corresponds to the initial MCMC samples before parameter convergence, the mean values of the remaining samples and the square root of the diagonal terms in their covariance matrix represent the plausible optimal values and uncertainties of the parameters, respectively.

3.3.2. Applied uncertainty propagation approach

As mentioned earlier, a forward UP analysis has been considered in this work to propagate uncertainties from thermodynamic parameters to Gibbs free energy of phases to microstructural characteristics in the $\text{Mg}_2\text{Si}_x\text{Sn}_{1-x}$ system. In this regard, a group of sampled parameter vectors has been considered as representative of the multi-variate posterior distribution of the thermodynamic parameters and then each selected sample (parameter vector) has been run through the CALPHAD model proposed in Section 3.1 to find the corresponding responses for Gibbs free energy of phases and the phase diagram. To find 95% Bayesian credible intervals

(BCI), 2.5% of the samples from both the lower and upper bounds of the response distributions have been discarded. The same approach has been used to propagate uncertainties from thermodynamic, microelastic and kinetic parameters to microstructural characteristics throughout a phase-field model. However, a Gaussian copula approach has been used to sample a reasonable number of parameter vectors from the distributions of microelastic and kinetic parameters—assuming they are independent—as well as the multi-variate posterior distribution of thermodynamic parameters obtained by the MCMC technique.

3.3.3. Sampling methodology

Here we note that propagating uncertainties from high-dimensional input sets has to be carried out in a way that minimizes the number of samples—a typical naive sampling scheme using MCMC approaches may require $\mathcal{O}(1,000,000)$ random samples—while at the same time accounting for the statistical correlations among input parameters—for example, parameters in CALPHAD thermodynamic descriptions tend to be highly correlated.

To construct sample sets with correct marginal distributions and preserved pairwise correlations, we instead make use of Gaussian copulas. A copula is a function that relates the joint cumulative distribution function (CDF) of multiple variables to their marginal CDFs and their correlations [79].

To begin, we assume that we have available marginal distributions, $f_{X_i}(x_i)$, for each parameter, where X_i denotes the random variable associated with the i th parameter, $i \in \{1, 2, \dots, K\}$ (K is the total number of parameters), and x_i is a particular realization of X_i . We also have available pairwise correlation coefficients, ρ_{ij} , where $i, j \in \{1, 2, \dots, K\}$, and

$$\rho_{i,j} = \frac{\text{Cov}(X_i, X_j)}{\sigma_{X_i}\sigma_{X_j}}, \quad (16)$$

where Cov denotes the covariance and σ denotes the standard deviation. These correlation coefficients are stored in a matrix, $R \in [-1, 1]^{K \times K}$, which we use to create the proper correlation structure among the pairwise joint distributions. To do this, we create a set of independent, identically distributed random vectors, $\mathbf{G}_1, \mathbf{G}_2, \dots, \mathbf{G}_n$ i.i.d. $\sim \mathcal{N}(\mathbf{0}, R)$, thus, each $\mathbf{G}_j$, $j \in \{1, 2, \dots, n\}$ is a K -dimensional random vector with a zero mean multivariate normal distribution with covariance, R . From this set of random vectors we can create a sample set, $\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_n$, of n samples from $\mathcal{N}(\mathbf{0}, R)$, where $\mathbf{g}_j$ are realizations from each identically distributed $\mathbf{G}_j$. Thus, we have n samples from standard normal distributions marginally, and each pairwise joint density has the desired correlation in the sample set.

From this set of samples, we generate uniformly distributed samples from $\{\mathbf{u}_j = (\Phi(g_j^1), \Phi(g_j^2), \dots, \Phi(g_j^K))\}_{j=1}^n$, where Φ is the cumulative distribution function of a standard normal random variable, and $\mathbf{g}_j = (g_j^1, g_j^2, \dots, g_j^K)$. This results in n samples from a vector of uniformly distributed marginal random variables with our desired pairwise correlations preserved. The final step makes use of the inverse cumulative marginal distributions of each of our parameters, $F_{X_i}^{-1}(u_j^i)$. From these inverse cumulative marginal distributions and the samples $\mathbf{u}_j = (u_j^1, u_j^2, \dots, u_j^K)$, we compute $\{\mathbf{x}_j = (F_{X_1}^{-1}(u_j^1), F_{X_2}^{-1}(u_j^2), \dots, F_{X_K}^{-1}(u_j^K))\}_{j=1}^n$, which is a set of n sample vectors sampled from the correct marginal distributions provided at the outset and preserved pairwise correlation information.

Our sampling methodology is demonstrated here for a two-dimensional random vector in Fig. 3 to make the preceding discussion more concrete. The top left plot of the figure is the original two-dimensional joint distribution with each marginal shown as well. This information is distilled into the marginal distributions,

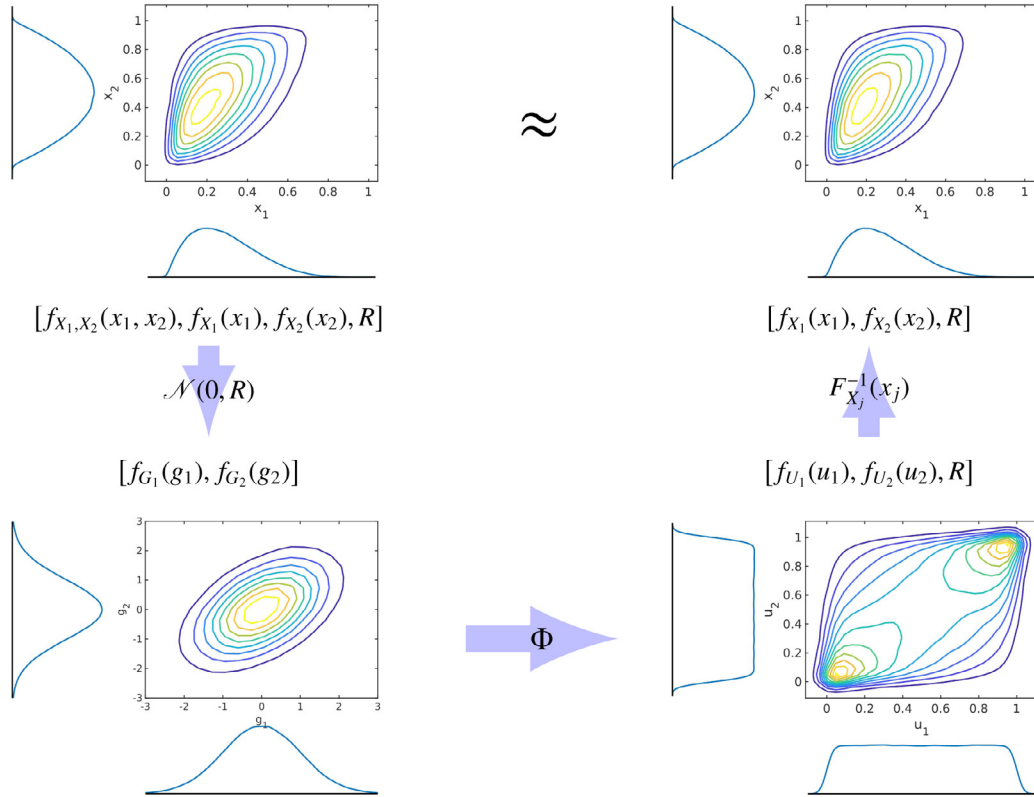


Fig. 3. Two-Dimensional demonstration of our sampling strategy that preserves marginal distributions and pairwise correlations. Updated figure.

$f_{X_1}(x_1)$ and $f_{X_2}(x_2)$, as well as the correlation coefficient matrix, R . From this correlation information, two correlated standard normal marginal distributions are created and shown in the bottom left plot of the figure. Samples from these distributions are passed individually through the standard normal cumulative distribution, Φ , which leads to two correlated uniform random variables shown in the bottom right plot. Samples from these distributions are passed individually through the inverse marginal distributions of each parameter, respectively. This results in samples from correct marginal distributions with preserved pairwise correlation. This can be seen by comparing the top left and top right plots. The top left plot is the true joint distribution and the top right plot is the joint distribution generated via this methodology, which is correct in terms of the marginal distributions and the correlation coefficient between X_1 and X_2 .

4. Results and discussion

Solid state reactions are not only influenced by chemical driving forces, but also by the interfacial and elastic energy contributions. The bulk free energy (f_{bulk}) greatly dominates the overall phase stability of a system but strong couplings to elastic fields arising from lattice and elastic mismatch as well as anisotropy in the properties play a dominant role in controlling the overall thermodynamic stability of the system as well as the topology and morphology of the resulting microstructures. Given the influence of bulk thermodynamic properties, the uncertainties in the bulk free energy require precise quantification first. This will be described in Section 4.1. Then the quantified uncertainties of the bulk free energy, the kinetic and microelastic parameters are propagated to microstructural characteristics in Section 4.2 through the elastochemical phase-field model.

4.1. Uncertainty quantification of Gibbs free energy parameters and phase diagram

A thorough parameter uncertainty analysis of the CALPHAD model for the $Mg_2Si_xSn_{1-x}$ pseudo-binary system has been performed through the MCMC technique (explained in Section 3.3.1) against the calculated composition-temperature data sampled from the phase diagram proposed by Kozlov et al. [63]. Then, the calculated parameter uncertainties have been propagated to the Gibbs free energy of phases and the phase diagram.

As shown in the Table 1, three parameters 0a , 0b and 1a are selected for each phase in the CALPHAD model, i.e. $Mg_2Si_xSn_{1-x}$ and liquid (six parameters in total). As mentioned earlier, the deterministically optimized parameters obtained from Thermo-Calc have been utilized as initial parameter values for MCMC sampling process; however, lack of knowledge about the parameter probability distributions resulted in the consideration of uniform (non-informative) prior distributions for the parameters. In addition, $\pm 50\%$ of the parameter initial values have been considered as the parameter ranges during this process.

During MCMC calibration, 100,000 samples have been generated to ensure parameter convergence. In this regard, plotting the joint frequency distribution of each pair of the model parameters can graphically show parameter convergence in parameter space. For example, one of these plots has been shown in Fig. 4. As observed in this figure, the red region with the highest density of parameter samples indicates the convergence region in the pair parameter space. Moreover, these kinds of plots can qualitatively show the degree of correlation between the model parameters based on the linearity and direction of the convergence region.

The degree of correlation between each parameter pair can also be quantified through the Pearson linear coefficient [80] denoted by ρ_{ij} in Eq. (16). It is worth noting that the linear coefficient is

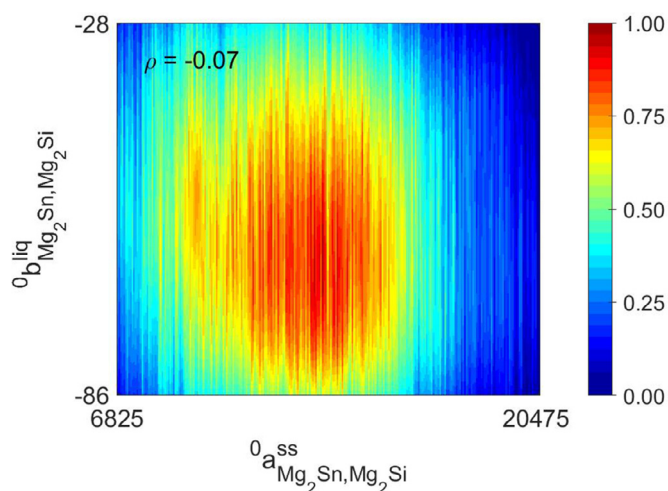


Fig. 4. Joint frequency distribution between a selected pair of parameters.

a quantity between -1 and 1 . Generally, parameters are uncorrelated if ρ is close to 0 , but highly correlated if ρ is close to either -1 or 1 . In addition, the negative and positive signs indicate the correlation direction. In Fig. 4, a semi-circular shape of convergence region and a very close value of ρ to zero imply negligible correlation between $0a_{ss}$ and $0b_{liq}$.

After discarding the burn-in period in the beginning of the MCMC sampling, the marginal probability distribution of each model parameter can be plotted as shown with blue color in Table 1 and Fig. 5. As observed in these distributions, all parameters show distributions close to truncated normal. Therefore, a truncated normal probability distribution has been fitted to each marginal probability distribution, where the hyper-parameters (μ and σ) of these truncated distributions have been estimated through a maximum likelihood method. The estimated μ and σ of each parameter probability distribution has also been listed in Table 1, which represent the plausible optimal values and uncertainties of the model parameters.

The parameter uncertainties in Table 1 have been propagated to the molar Gibbs free energy of mixing for different phases in the system to the phase diagram. As mentioned in Section 3.3.2, the last 5000 MCMC samples as an ensemble of the whole convergence region have been used in a forward model analysis scheme to identify the variation of the phase boundaries and invariant lines in the phase diagram. 95 BCIs have then been determined by discarding 2.5% of the resulting samples from above and below the variation band at any specified Si composition in the Gibbs energy of mixing curves or the phase diagram. Fig. 5 demonstrates uncertainty propagation across different levels of CALPHAD modeling. Here, just an example of molar Gibbs energy of mixing for solid phases ($Mg_2Si_xSn_{1-x}$) at 700 K has been plotted to show how uncertainties propagate from the Gibbs free energy curves to the phase diagram. This analysis has been repeated for different temperatures in the range from 200 to 1400 K to construct the whole phase diagram and its uncertainties. Based on Fig. 5, it should be noted that there are very high uncertainties in the curves for the Gibbs energy of mixing as well as the resulting phase diagram for most temperatures. The considerable uncertainties in the Gibbs energy curves are a direct consequence of the large uncertainty in the location of the phase boundaries in this system.

4.2. Effect of parameter uncertainties on microstructure

A complete phase-field model for the $Mg_2Si_xSn_{1-x}$ system requires the definition of a large number of parameters—see Table 1.

The proper characterization of all relevant parameters has been a major challenge in mesoscale microstructural models, and it is a common practice to define the values of these parameters deterministically and ignore their actual variabilities/uncertainties. As a consequence, exclusively deterministic models completely sidestep the intrinsic uncertainty in model outputs. This implies that common sources of uncertainty may have little or no impact on the selection of the value of model parameters. Here, the uncertainties in parameters of the Gibbs energy, obtained through the MCMC approach, are combined with the prior distribution of the parameters of other sub-models (microelasticity and kinetic) to study the microstructural evolution using the phase-field model. The prior information of the all parameters are reported in Table 1. In the following, we elaborate on the source of these prior distributions.

The process of phase separation is influenced by the elastic anisotropy, and the hardness enhancement observed upon the age hardening relies on a shear modulus difference between the formed domains as well as their coherency strain [81]. Thus, it is of primary interest to take into account the uncertainties in the elastic properties of the $Mg_2Si_xSn_{1-x}$ system to study their contributions to the variability/uncertainty in microstructural predictions.

The *ab-initio*-based elastic constants (C_{11} , C_{12} , and C_{44}) for (c)- Mg_2Sn and (c)- Mg_2Si are reported in the literature—we note that these quantities have yet to be measured experimentally. These include both 0 K, and high temperature data which are calculated by taking into account entropic and/or quasi-harmonic effects on the free energies of the lattices. The CDF of C_{11} , C_{12} , and C_{44} for both phases are shown in the left side of the microstructure palette in Fig. 6(a). Both *ab-initio* and experimental lattice parameters for the cubic $Mg_2Si_xSn_{1-x}$ system at room/high temperature for different ranges of compositions are considered. These values are either provided for individual phases (Mg_2Sn and Mg_2Si), or for the parent phase as a function of composition. The CDF of the lattice parameter data are shown in Fig. 6(b). Using these data, the range of SFTS (ϵ_{ij}^0) for Mg_2Sn and Mg_2Si is estimated. This range is used to draw samples from a uniform distribution in order to propagate the uncertainty in the microelasticity model.

The molar volumes (V_m) of the phases are also obtained based on both calculated and measured data. The calculated data is obtained from $V_m^{cryst} = V_{cell} \cdot N_v / Z$ where V_{cell} is the volume of the unit cell, N_v is Avogadro's number, and Z is the number of atoms in the cell. $Z = 3$ for both phases considering the fact that the crystal structure of both Mg_2Sn and Mg_2Si is $Fm\bar{3}m$. The distribution is shown in Fig. 6(c), which is very similar to the distribution of lattice parameters shown in this figure. The lower bound of this distribution belongs to the data obtained by division of the molar mass and density distributions.

Experimental diffusion data are used to estimate the atomic mobility. The recent combinatorial diffusion couple study by Vivés et al. [62], and the indirect calculation of inter-diffusion coefficient using a forward-simulation method provides an insight about the diffusion at 600 °C. Their calculations suggest an estimated diffusion coefficient of $\approx 10^{-15} \text{ ms}^{-2}$. There is no other information regarding diffusivity measurements to the best of our knowledge. Accordingly, the mobility is estimated by the Einstein's rule ($M = \frac{D}{RT}$), or ($M = \frac{D}{(\frac{\partial^2 f}{\partial c^2})}$) [40]. Hence, a uniform sampling is carried out around this value. Initial composition of the alloy is taken to be uniformly distributed with c_{Mg_2Si} between $0.3 - 0.5$ to make sure we are under the chemical miscibility region.

In order to propagate the uncertainties of the prior data, it is necessary to carry out high-throughput phase-field simulations of microstructure evolution in the $Mg_2Si_xSn_{1-x}$ system. Using the strategy described in Section 3.3.3, we sampled 10,000 combinations of the parameters out of the prior distributions, which have

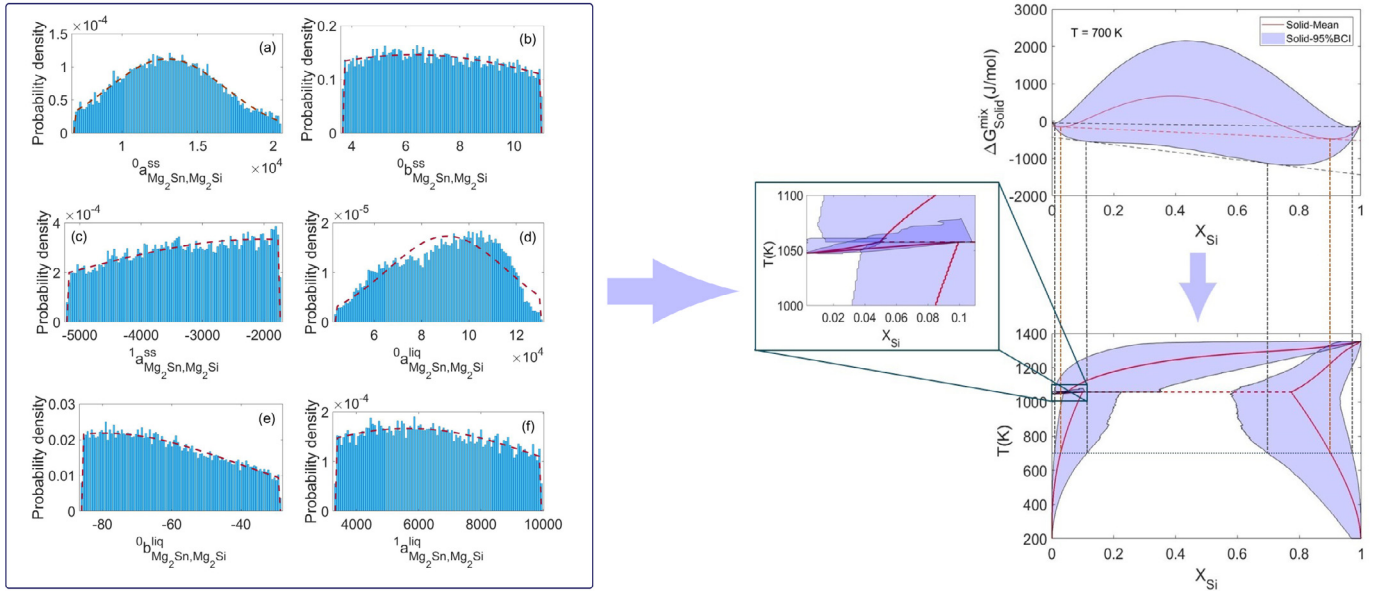


Fig. 5. An illustration of UP from thermodynamic parameters to the molar Gibbs energy of mixing, to the phase diagram for $\text{Mg}_2\text{Si}_x\text{Sn}_{1-x}$ quasi-binary system at 700 K. left) the probability distributions for the six CALPHAD parameters, right) the calculated Gibbs energy of mixing at 700 K and uncertainty bands and the pseudo-binary phase diagram with a close-up of the peritectic reaction. Updated Figure.

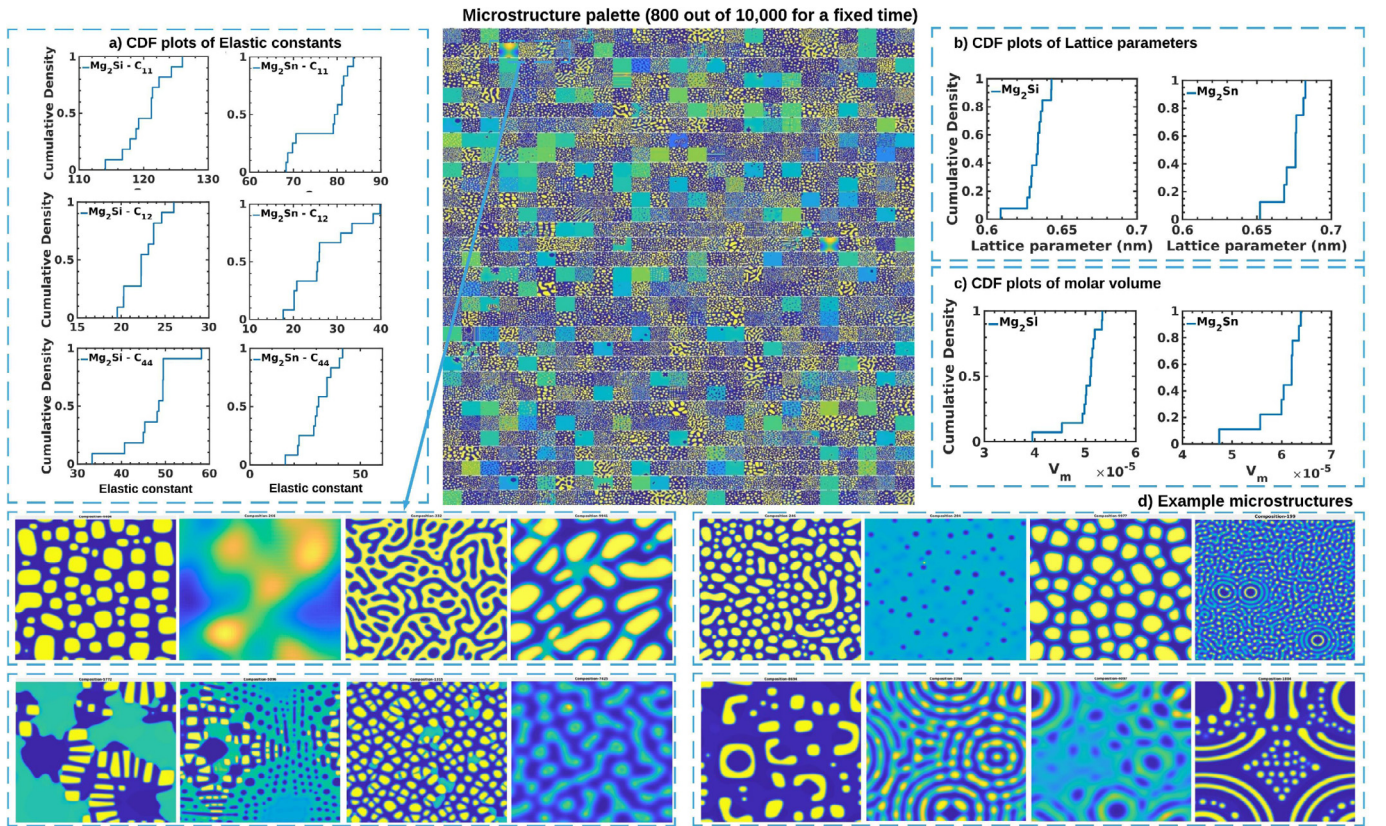


Fig. 6. (Center) Microstructure library from 800 phase-field runs (out of 10,000 run) for the $\text{Mg}_2\text{Si}_x\text{Sn}_{1-x}$ system at a fixed time. 10,000 points in a 18-dimensional parameter space were sampled from the prior distributions shown in Table 1. CDF plots for (a) elastic constants, (b) lattice parameters, (c) molar volumes of the two product phases. (d) Example microstructures extracted from the palette. Updated figure.

been fed to the phase-field equation solver as implemented in a Fortran program. The whole process of 1) data initialization, 2) environment preparation and 3) Fortran code run in the Terra supercomputing cluster at Texas A&M University is automated through a Python wrapper. This is a custom batch job submission routine that we developed and it provides a convenient way to run large num-

ber of serial or multi-threaded jobs in the supercomputer. A square simulation cell with 512×512 grid points, where $L_x = L_y = 350$ nm is used to perform the simulations. All simulations are performed at isothermal temperature of 720°C . The snapshots of the obtained microstructures for 800 samples are shown in the form of a palette in Fig. 6. Sixteen example microstructures (at the same physical

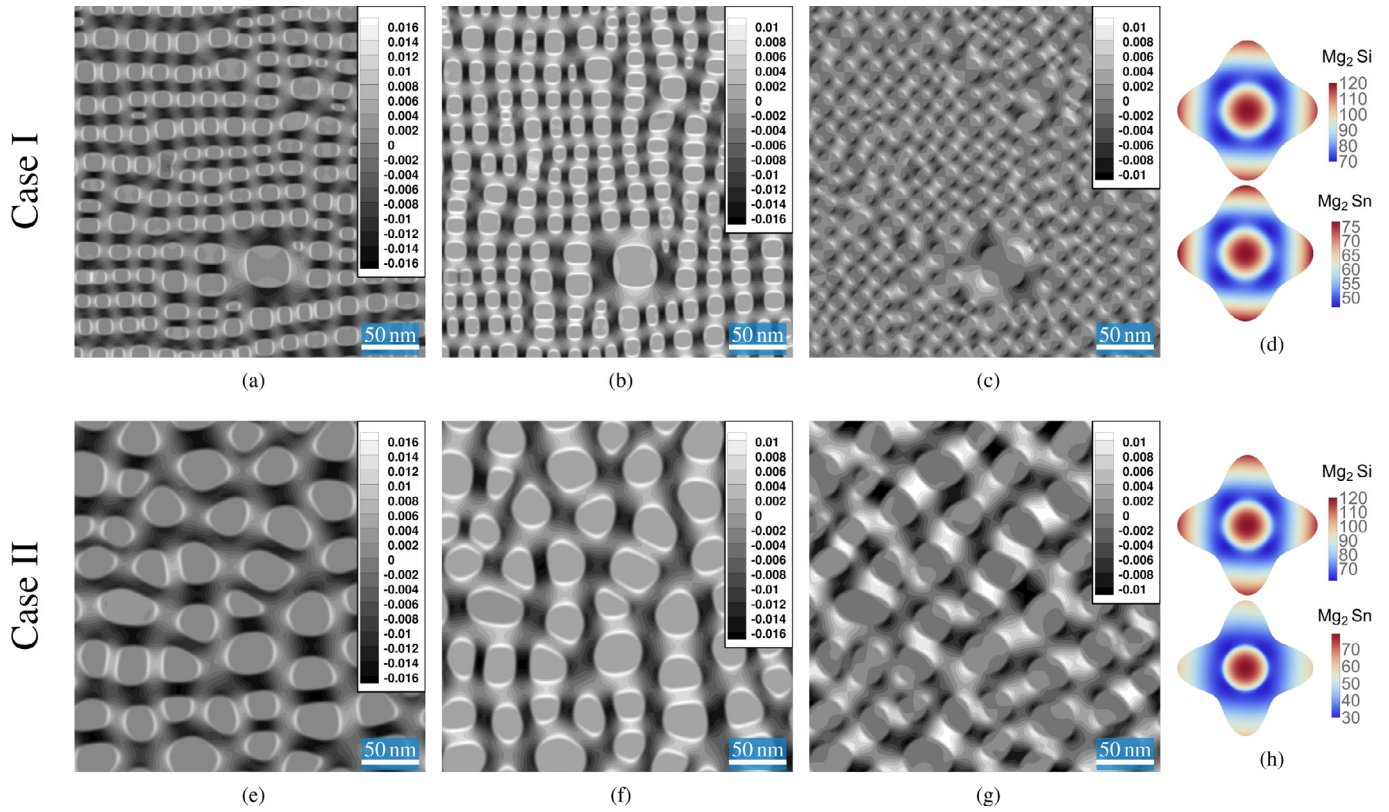


Fig. 7. Effect of the local strain due to inhomogeneous elastic effects in the microstructure. Although, the parameters are very close in the two cases, the final microstructure is different due to slight change in C_{44} component of the elastic tensor of Mg_2Sn . (a) and (e) ε_{11}^{el} , (b) and (f) ε_{22}^{el} , (c) and (g) ε_{12}^{el} , and, (d) and (h) the directional dependence of elastic tensors of Mg_2Si and Mg_2Sn phases in 2D space. Alloy composition and elastic parameters are summarized here. Case I: $c = 0.473$, the SFTS value is $\varepsilon_{ii}^0 = -0.0176$, elastic constants are $[\text{Mg}_2\text{Si}: C_{11} = 120.2, C_{12} = 22.1, C_{44} = 56.3 \text{ GPa}]$, $[\text{Mg}_2\text{Sn}: C_{11} = 76.9, C_{12} = 19.2, C_{44} = 38.4 \text{ GPa}]$. Case II: $c = 0.495$, $\varepsilon_{ii}^0 = -0.0191$, $[\text{Mg}_2\text{Si}: C_{11} = 120.5, C_{12} = 22.2, C_{44} = 48.8 \text{ GPa}]$, $[\text{Mg}_2\text{Sn}: C_{11} = 79.14, C_{12} = 38.2, C_{44} = 21.5 \text{ GPa}]$. For interpretation of the colors refer to the online version of this document. Updated figure.

evolution time) are also extracted from the palette for more effective demonstration and shown at the bottom section of this figure.

Phase-field models tend to be highly nonlinear so that its output can differ qualitatively, depending on the region in the input/parameter space where the sample is taken from. Phase-field models are also highly complex in their formulation and are thus not amenable to intrusive approaches to UP. Moreover, they are computationally expensive, with full three-dimensional realizations of the simulations requiring upwards of 10,000 CPU-hours in some of the fastest supercomputers available. To reduce the modeling cost, we only performed 2D calculations in this study. Finally, the input/parameter space is high-dimensional, with more than 20 stochastic input conditions and model parameters that dramatically affect the thermodynamic and kinetic state of the system, as shown in Fig. 6.

It must be noted that since we are interested in the elastochemical interactions in the microstructure evolution of the $\text{Mg}_2\text{Si}_x\text{Sn}_{1-x}$ system, the effect of variations in elastic constants on the microstructure were taken into account. The local long-range interactions (i.e. strain and/or stress fields) of the heterogeneous multi-phase nanostructure are quite sensitive to the selection of the elastic parameters, and therefore the latter must be considered with a great care. Fig. 7 shows the elastic strain maps (ε_{ij}^{el}) for two distinct microstructures that are obtained with different sets of elastic constants (i.e. shear constants), and very similar eigenstrain (ε_{ii}^0) values. Both cases might be categorized as cuboid-type microstructures. In the first case, (Fig. 7a–c) the morphology is perfectly cuboidal with a slight deviation from unimodal particle-size distribution. In the second case (Fig. 7e–g), the particles are

sheared at the corners and evolve towards a cuboidal shape during the coarsening stages. The elastic surface plots in the right hand side of each microstructure case shown in Fig. 7 demonstrate the elastic tensor surface plots as a function of direction for each of the product phases.

4.3. Microstructure quantification

Uncertainty propagation consists of linking the uncertainty in the input conditions and model parameters to the generated uncertainty in the model output. To properly quantify the uncertainty and study the way it propagates across the simulations, it is necessary to define QoIs whose distributions can then be interpreted as a measure of variance resulting from the stochastic nature of the input space. In the case of microstructures, these QoIs could correspond to different metrics connected to the morphology, topology and overall distribution of microstructural features.

A challenging aspect of the uncertainty propagation effort in this work is the large dimensionality of both the input and output spaces. The large dimensionality of the input space originates from the many parameters that are necessary to complete the phase-field model. Moreover, given the fact that the output microstructure space is highly heterogeneous, it is to be expected that many QoIs would be necessary to completely characterize the effect of model inputs on the morphology and topology of the resulting microstructures. In order to handle these very complex spaces, we will resort to machine learning frameworks, as will be described below.

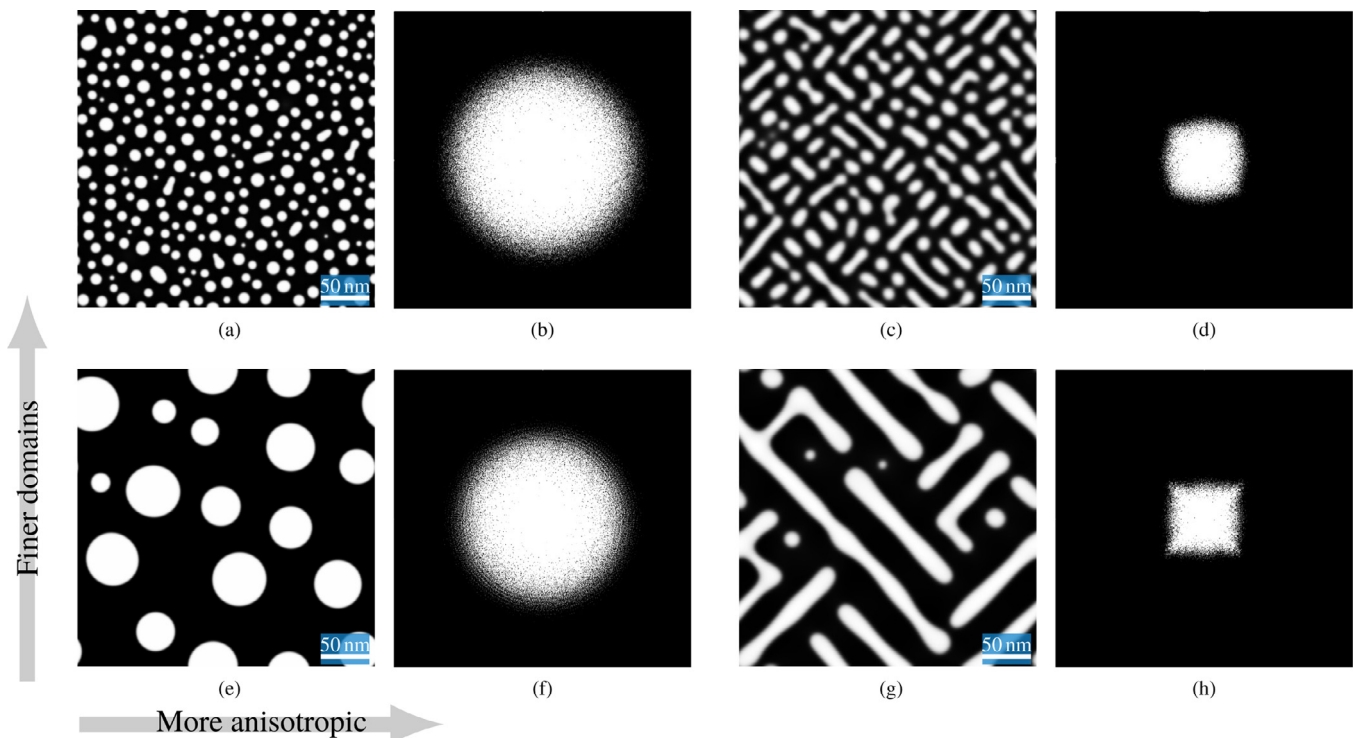


Fig. 8. (a,c,e,g) Example microstructures (composition field, c), and the corresponding (b,d,f,h) power spectrum maps, ($\log(|c|_k^2)$) converted to gray scale and shifted to the center obtained by Fourier transformation of the microstructures. The finer the domains, the power spectrum circle shape becomes larger. In turn, the anisotropy changes the shape from circle to square. Updated figure.

The propagation of uncertainty first requires the quantification of changes in the microstructure. Given the large dimensionality of the model output (512×512 grid points), dimensional reduction is necessary. Moreover, in order to enhance the interpretability of the analysis it is desirable for the reduced dimensions to have physical significance. The determination of QoIs typically involves extraction of features and identifying a bank of descriptors that can be used to train a classifier based on the frequency of observations [82,83]. A conventional yet very useful approach is to use semantic texton forests [84]. This is specifically useful when the phase-field variable is composition and can also be broadly used in the real-image data. In addition, visual words [85] can be used as powerful discrete image representations for categorization. Another way to tackle this is to utilize one or a combination of Filter-bank responses (e.g. Fourier or other sort of wavelets), and one or combination of invariant descriptors (e.g. SIFT [86,87]).

A series of microstructure analytic tools were used to determine a wide range of QoIs to evaluate the developed UP framework. We use metrics such as average feature size, area fraction, composition of the phases, aspect ratio as well as increasingly common approaches such as n-point statistics [83], and frequency-domain analysis (cf. Fig. 8). In latter case, the general information about the morphology and orientation of the particles are reflected in the frequency space. Many of the commonly used metrics provide information about the average state of microstructure spaces, but in many cases they do not provide information about their topology. In the case of transport behavior (such as phonon conductivity), topology may play an important role and metrics quantifying this microstructure feature are necessary.

Table 2 summarizes the list of QoIs and the associated statistics for phase-field model outputs. It contains the posterior, means, and index of dispersion denoted by variance to mean ratio (VMR) of the obtained QoI. VMR is an index that quantifies the dispersion of a probability distribution and measures the clusterabil-

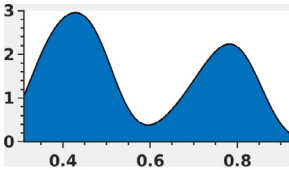
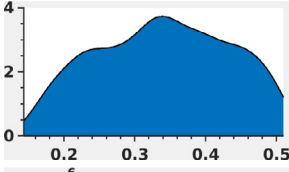
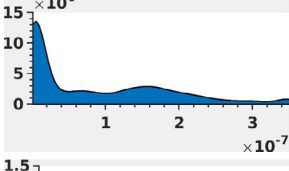
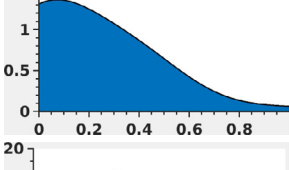
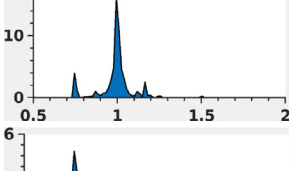
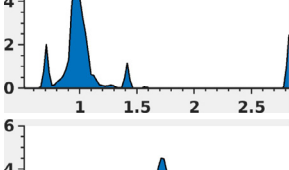
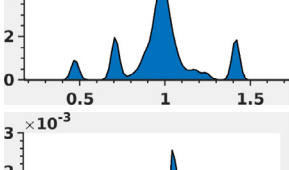
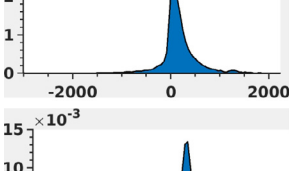
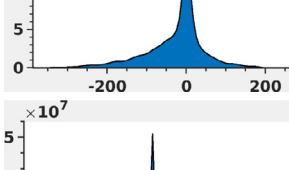
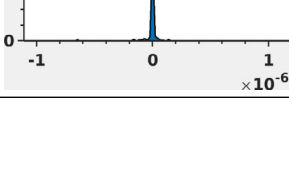
ity and variability in the data. When $\text{VMR} > 1$, the data is over-dispersed, and when $0 < \text{VMR} < 1$, the data is under-dispersed. $\text{VMR} = 0$ corresponds to a random data-set and $\text{VMR} = 1$ corresponds to a Poisson distribution. Except for the last three QoIs, all other QoIs show multimodal distributions. We wish to note that this multi-modality would render simplified uncertainty analysis frameworks—such as those based on sampling of minimum, maximum and mean values of input parameters—ineffective.

It should be noted that $c_{\text{Mg}_2\text{Si}}$ and $c_{\text{Mg}_2\text{Sn}}$ are considered to be the equilibrium compositions—in molar fractions—of the Mg_2Si and Mg_2Sn phases, respectively. $c_{\text{Mg}_2\text{Si}}$ posterior clearly illustrates a bimodal distribution, while $c_{\text{Mg}_2\text{Sn}}$ illustrates a trimodal distribution. In the $c_{\text{Mg}_2\text{Si}}$ posterior, the first (left) peak belongs to the case where the process of phase separation has not finished and the microstructure remains non-decomposed. In principle, this is due to either slow kinetics or tendency of the system to dissolve rather than phase separation. In the latter, elastic driving force overcomes the chemical driving force. In this case, the elastic driving forces are significantly stronger than the chemical one. This is believed to be achievable in practice through non-equilibrium synthesis conditions—refer to [40] for further information. Therefore, the non-decomposed microstructures may be classified to two sub-category. In the $c_{\text{Mg}_2\text{Sn}}$ posterior, one major peak and two minor close-by peaks are present. Similar to $c_{\text{Mg}_2\text{Si}}$ posterior, these peaks belong to decomposed and non-decomposed categories. The right peak belongs to decomposed microstructures. On the contrary to $c_{\text{Mg}_2\text{Si}}$ where the two subcategories of non-decomposed microstructures are not directly distinguishable, here two separate peaks can be seen for the two subcategories.

The characteristic length scale for each microstructure is obtained by calculating the radial basis power spectrum of the Fourier transform of the microstructure (2D composition space). The distribution is bimodal, although highly skewed towards the low end of the range and its peaks belong to either the non-

Table 2

List of extracted QoI and their posterior distributions, mean, standard deviation, minimum and maximum from the entire set of phase-field runs.

Target variables	Posterior Distribution	μ	Index of Dispersion ($\frac{\sigma^2}{\mu}$)
c_{Mg_2Si} [mol]		0.57	0.056
c_{Mg_2Sn} [mol]		0.34	0.026
Characteristic Length [m]		8.62×10^{-8}	1.08×10^{-7}
Area fraction		0.17	0.30
Roundness		0.99	0.70
Diagonal/width		1.18	0.29
Diagonal/height		0.98	0.05
$\bar{\mu}_{chem}$ [Jmol ⁻¹]		-156.06	-1012.1
$\bar{\mu}_{elas}$ [Jmol ⁻¹]		-27.83	-214.75
$\bar{\mu}_{int}$ [Jmol ⁻¹]		1.59×10^{-9}	-6.01×10^{-6}

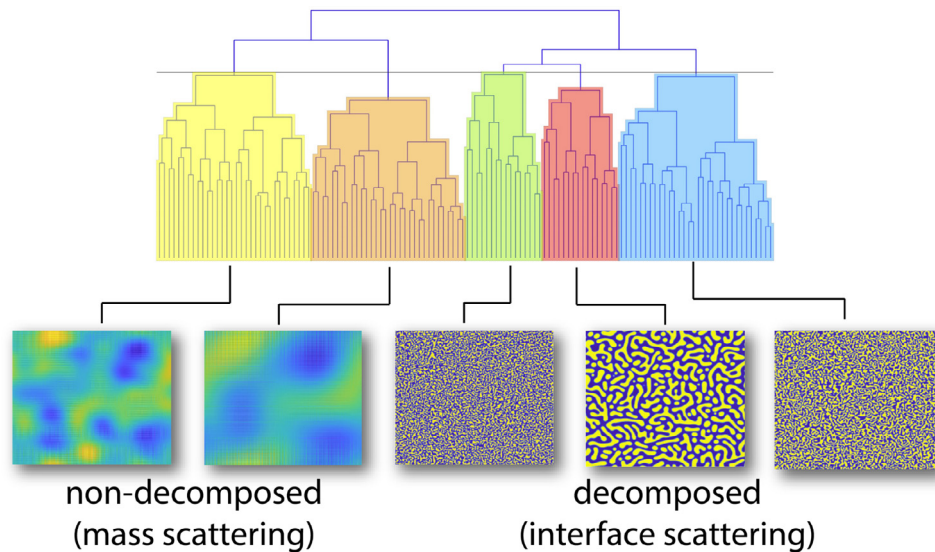


Fig. 9. Hierarchical clustering using Mahalanobis distance matrix. The horizontal line indicates 78% height ratio. The images are clustered into two main subcategories of 'decomposed' versus 'non-decomposed' microstructures. Each of these main categories are also divided into other sub-levels. Updated figure.

decomposed, or the decomposed structures with certain morphologies. The calculated area fractions uniformly ascend from a zero fraction for the dissolved or non-decomposed microstructures to the decomposed ones. The probability density of observation of non-decomposed or dissolved microstructures is higher than its counterpart for the decomposed ones. Roundness is extracted based on the shape of the morphology that appears in the power spectrum domain. Fig. 8 demonstrates examples of different power spectra that are obtained for the shown microstructures. For microstructures with somewhat isotropic morphologies the power spectrum is symmetric and appears to be a circle. On the other hand, the power spectrum appears to be in the form of rectangle/square for a perfectly oriented, anisotropic, morphology in which elastic effects tend to dominate the microstructure evolution. The shapes of the power spectrum become more defined as the microstructure becomes more coarse.

Quantities $Diagonal/width$ and $Diagonal/height$ are the relative ratio of the diagonals of the power spectrum centroid with respect to its width and height. The closer the value is to $\sqrt{2 \cdot height^2}$ and/or $\sqrt{2 \cdot width^2}$, the morphologies tend to shear more. Moreover, the higher these values get, the anisotropy becomes more evident. On the other hand, the mean value of each of the contributing energy fields are also obtained for the current sampling. The mean of bulk and elastic driving forces in the domain exhibit unimodal probability distributions. However, the mean of interfacial driving forces demonstrates multimodal isolated peaks.

Table 2 represents the total variance in the microstructure space assuming the CDF of the input parameters and their statistical correlation. The propagated uncertainty is represented through different QoIs. However, these results are aggregated and further analysis is necessary in order to examine the effect of individual parameters (or sets of parameters) on the resulting microstructures. The large dimensionality of the input and output spaces, as we discussed above, makes it necessary to rely on machine learning approaches that facilitate the analysis of the microstructure space and their relation to model inputs.

4.4. Application of the materials informatics techniques in microstructural evaluation

A primary purpose of data-mining techniques is to facilitate the determination of possible meaningful patterns in a given space

in order to better understand, for example, the connection between inputs (or features) and the QoIs, as well as to make informed decisions based on the patterns (or models) elicited from the data under analysis [88]. Materials informatics is an interdisciplinary blend of statistics, machine learning, artificial intelligence, pattern recognition and materials science. Here we employ a few core tasks (e.g. cluster analysis, classification, etc.) on the posterior data generated in this study. This collective integration of statistical learning tools with experimental and/or computational materials science allows for an informatics-driven strategy for materials analysis design.

A preliminary simple hierarchical clustering analysis, using different distance metrics, of the microstructural set produced from the uncertainty propagation exercise is performed to elucidate the structure of the available microstructural data. The results are demonstrated as a corresponding dendrogram in Fig. 9. This calculation is based on the 18 material parameter as the inputs and the eight QoIs as the target variables. In addition, several distance metrics (i.e. Euclidean, Manhattan, Mahalanobis, Spearman and Pearson) are used. The analysis shown in Fig. 9 is based on Mahalanobis distance metric where the elements are taken in a pairwise fashion in the given set using $d_M(x, y) = \sqrt{(x - y)^T S^{-1} (x - y)}$ where S is the covariance matrix. The dendrogram shown in Fig. 9 is created by the Ward's linkage criterion (increase in variance for the cluster being merged) and aims to indicate the similarity/dissimilarity among annotation categories. The five sub-clusters of the dendrogram shown in different colors are annotated with selected representative microstructures. As seen, this dendrogram clusters the microstructures into two main categories, i.e. 'decomposed' and 'non-decomposed' classes, with some error. The quantification of the error requires human interaction and manual identification of the outliers. Instead, we currently work toward unsupervised classification of the microstructure images.

A comparison of several classifiers is performed on the extracted QoIs data-set. Classification methods are suited for cases where the class label is discrete. Hence, we simply use the 'non-decomposed' versus 'decomposed' class labels. Here we return to the connection with performance in TE materials by noting that microstructures that correspond to non decomposed states could be associated with alloying/mass phonon scattering, while decomposed microstructures corresponds to interfacial phonon scattering. The length scales of different scattering mechanisms are

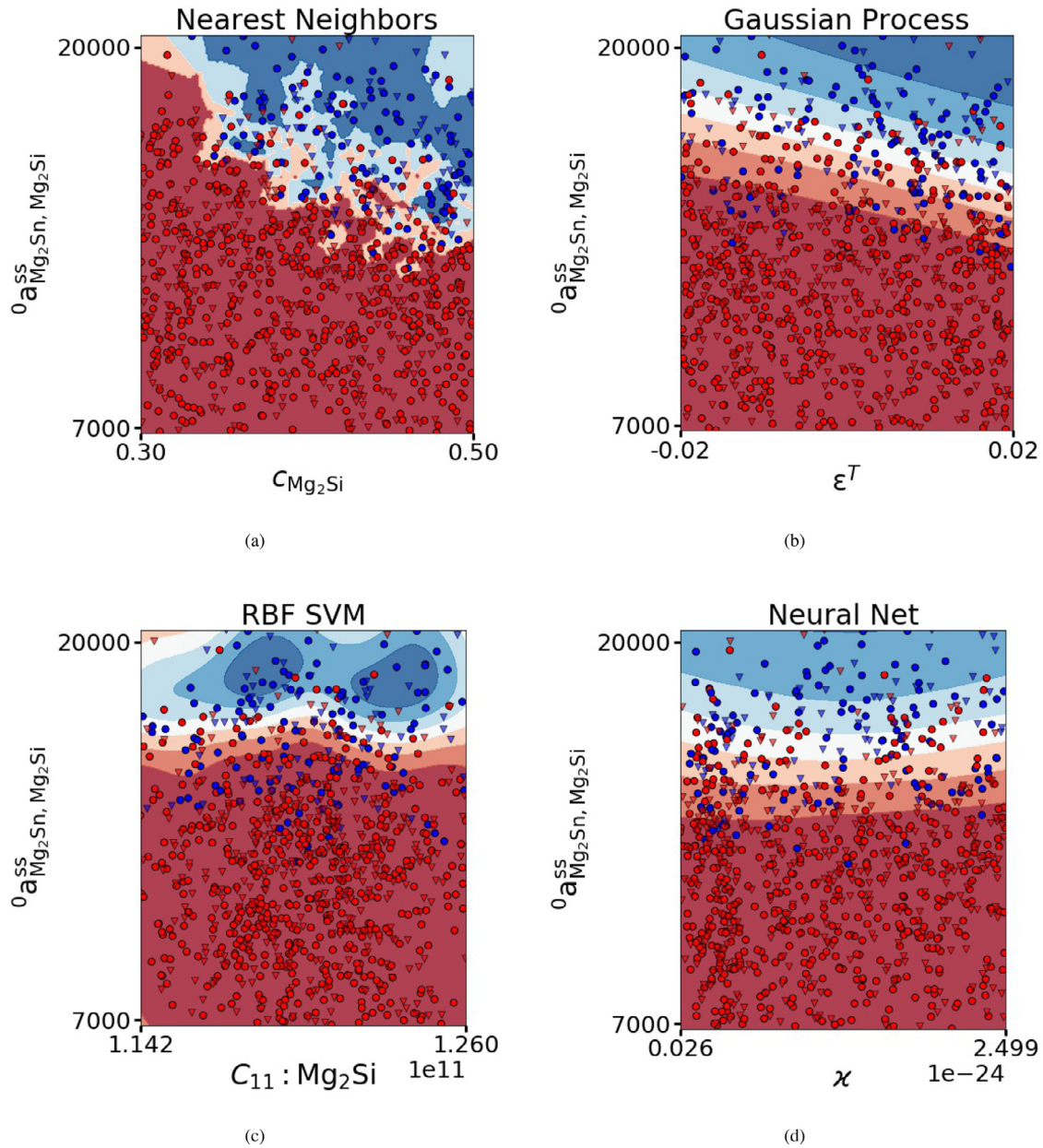


Fig. 10. Classification results for (a) Nearest neighbors classifier in (alloy composition ($C_{\text{Mg}_2\text{Si}}$), $0a_{\text{Mg}_2\text{Sn}, \text{Mg}_2\text{Si}}^{\text{ss}}$) space, (b) Gaussian process classifier in (ϵ^T , $0a_{\text{Mg}_2\text{Sn}, \text{Mg}_2\text{Si}}^{\text{ss}}$) space, (c) Radial basis function support vector machine (RBF SVM) classifier in ($C_{11} : \text{Mg}_2\text{Si}$, $0a_{\text{Mg}_2\text{Sn}, \text{Mg}_2\text{Si}}^{\text{ss}}$) space (d) Neural network classifier in (χ , $0a_{\text{Mg}_2\text{Sn}, \text{Mg}_2\text{Si}}^{\text{ss}}$) space. ■ stands for non-decomposed or dissolved microstructures and ■ stands for decomposed microstructures with different morphologies. Circle markers represent the training set and the triangles represent the test set. For interpretation of the colors the reader is referred to the online version of this document.

different and are thus expected to change the phonon transport characteristics and the corresponding thermoelectric performance of the $\text{Mg}_2\text{Si}_x\text{Sn}_{1-x}$ system [57].

Fig. 10 illustrates the nature of such decision boundaries in the 2D input parameter regions which can be used to determine the desired regions (alloy/mass vs. interface scattering) in the material parameter space. Several classifiers are tested on all of the parameter pairs and the selected classifiers to report are the Nearest neighbors, Gaussian process, Radial Basis Function (RBF) kernel Support Vector Machine (SVM), and Neural Network. Each subfigure in Fig. 10 illustrates the nature of decision boundaries for a given pair of the input parameter using different classifiers. The training points in these plots are shown in solid colors and the testing points are defined by the semi-transparent contour. We

first randomly extracted 10% of the original data obtained from a set of microstructures frozen at a fixed time. 60% of this data is used as the training set and 40% as the test set. Fig. 10a demonstrates the nearest neighbors classifier in the (C , $0a_{\text{Mg}_2\text{Sn}, \text{Mg}_2\text{Si}}^{\text{ss}}$) space with a transitioning boundary where the points are mixed in the boundary of the two class. Fig. 10b demonstrates the Gaussian process classifier [89] result in the (ϵ^T , $0a_{\text{Mg}_2\text{Sn}, \text{Mg}_2\text{Si}}^{\text{ss}}$) space with a smooth, and continuous separation boundary between the decomposed (blue) and non-decomposed (red) clusters. Fig. 10c shows the classification result in ($C_{11} : \text{Mg}_2\text{Si}$, $0a_{\text{Mg}_2\text{Sn}, \text{Mg}_2\text{Si}}^{\text{ss}}$) space using the RBF SVM classifier that reveals the transitioning boundaries between the two clusters. The red cluster preserves a larger area while certain points of this cluster are mixed with the blue clus-

ter. The result associated with the trained neural network classifier in $(x, 0q_{\text{Mg}_2\text{Sn}, \text{Mg}_2\text{Si}}^{\text{ss}})$ space is shown in Fig. 10d, which again illustrates a smooth, and continuous separation boundary between the two classes. However, it should be noted that no clear classification boundaries are obtained in most pair-parameter spaces which can be attributed to the lack of recognition through 2D projections of the high dimensional parameter space. We note that in these cases, the application of a dimensional reduction technique, e.g. principal component analysis (PCA), might be helpful for more rigorous classification study.

While further investigation of the microstructure space through machine learning approaches may be warranted in order to fully characterize the connections between input parameters and resulting microstructures, this is well beyond the scope of the present work. It is in fact our intent to further explore the use of different machine learning frameworks with the obtained microstructure dataset.

5. Summary and conclusion

As with any computational analysis, uncertainty quantification/propagation plays a major role in predicting the outcome of multi-scale models in materials science. One of the most important tasks in materials design under the ICME framework is propagation of uncertainties of parameters across the multi-scale models that connect process-structure-property-performance. Hence, we addressed an existing challenge in materials science, i.e., UP from the thermodynamic parameters to the microstructural features through a chain of CALPHAD, microelasticity, and phase-field models. This is a multi-step uncertainty propagation with the techniques used in each step motivated by the type of problem that is being dealt with. Technically, thermodynamic parameters in the CALPHAD model and their underlying imposed uncertainties obtained from an MCMC sampling approach has been propagated to Gibbs free energy of phases and equilibrium phase diagram through a forward analysis of an ensemble of these samples. Then, these uncertainties as well as the uncertainties of microelastic and kinetic parameters have been propagated to the microstructural features using a Gaussian copula sampling approach.

One of the most important aims of the present study is to rectify the common belief towards the deterministic assumption about the parameter values in the phase-field models. In other words, the variations of the parameter values due to their uncertainties can result in a range of microstructures rather than just a single microstructure obtained from the deterministic view. The propagation of uncertainty in the prior parameter space using model chains resulted in a massive microstructure data-set (~ 50 TB) that is curated in the OPMD website [90]. The quantities of interest in the microstructures are identified and extracted using automated frameworks from the large amounts of data that is generated by HT-phase-field runs. We used eight QoIs to map the obtained probability distributions of parameters into probability distributions of the extracted quantities. Then, data-mining techniques are employed to find patterns in the parameter space that can contribute to better understanding of process-microstructure relations. The results show that the data is clustered into two main categories of ‘decomposed’ and ‘non-decomposed’ microstructures. These results can be very useful for engineering material behavior in favor of specific phonon scattering mechanism and/or better thermoelectric response. The proposed framework is generalizable to applications to other materials problems and microstructure-sensitive properties. As of this writing, further exploration of the input parameter-microstructure space is in progress and will be reported in our future work.

Developing a general QoI that represents the microstructural space broadly will enable better quantification of the uncer-

tainty propagated through the models and helps establishing inverse maps to connect regions in the microstructure space to the corresponding regions in the input space. Standard methods of microstructure analysis are not good enough to represent microstructure spaces in a general way. To establish process-structure-property connections, it is very important to find a relationship between the specific features of such microstructures and their properties.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors would like to acknowledge the Terra supercomputing facility of the Texas A&M University, for providing the computing resources to carry out the research reported in this paper. The data is currently hosted by this facility. This research was supported by the National Science Foundation under NSF grant no. CMMI-1462255 and CMMI-1663130. RA and DA also acknowledge the support of ARL through [grant no. W911NF-132-0018] as well as NSF through grant no. DGE-1545403.

Appendix A. Semi-implicit algorithm for solving C-H Equation

The semi-implicit algorithm for solving the C-H equation is adopted from Ref. [72]. The fourth-order C-H equation, neglecting the higher-order concentration gradients has the following form:

$$\frac{\partial c}{\partial t} = -\nabla^2 \left(\frac{\partial f^0}{\partial c} + \mu^{\text{elastic}} + \nabla^2 c \right) \quad (\text{A.1})$$

where f^0 is the free energy of homogeneous domains in the material, and μ^{elastic} is the elastic potential for phase change obtained by means of the microelasticity model. The Fourier transform of this equation is:

$$\frac{\partial \{c\}_{\mathbf{k}}}{\partial t} = -k^2 \left[\left\{ \frac{\partial f^0}{\partial c} \right\}_{\mathbf{k}} + \{\mu^{\text{elastic}}\}_{\mathbf{k}} + 2k^2 \{c\}_{\mathbf{k}} \right] \quad (\text{A.2})$$

where $\mathbf{k}$ is a vector in the reciprocal space, and $k = |\mathbf{k}|$. The semi-implicit discretization of this equation is:

$$\frac{c(\mathbf{k}, t + \Delta t) - c(\mathbf{k}, t)}{\Delta t} = -k^2 \left\{ \frac{\partial f^0}{\partial c} \right\}_{\mathbf{k}} - k^2 \{\mu^{\text{el}}\}_{\mathbf{k}} - 2k^4 c(\mathbf{k}, t + \Delta t) \quad (\text{A.3})$$

where Δt is the time step for the numerical integration. The Fortran compiler and (discrete) Fourier transforms needed for our calculations have been carried out using Intel(R) Fortran Compiler for Linux and Intel(R) Math Kernel Library for Linux, respectively.

Appendix B. Comparison of experimental and synthetic microstructures

In this section, we provide several example microstructures to illustrate the diversity of the microstructures in the Open Phase-field Microstructure Database with reference to the experimental cases. While all calculations are performed for $\text{Mg}_2\text{Si}_x\text{Sn}_{1-x}$ system, the propagation of uncertainty in multiscale chain of the models resulted in the extensive range of microstructures. Fig. B.1 summarizes a series of these experimental and computational microstructures.

A typical scanning electron micrograph (SEM) of the $\gamma + \gamma'$ structure of Co-9.2Al-10.2W(at.%) is shown in Fig. B.1(a), where

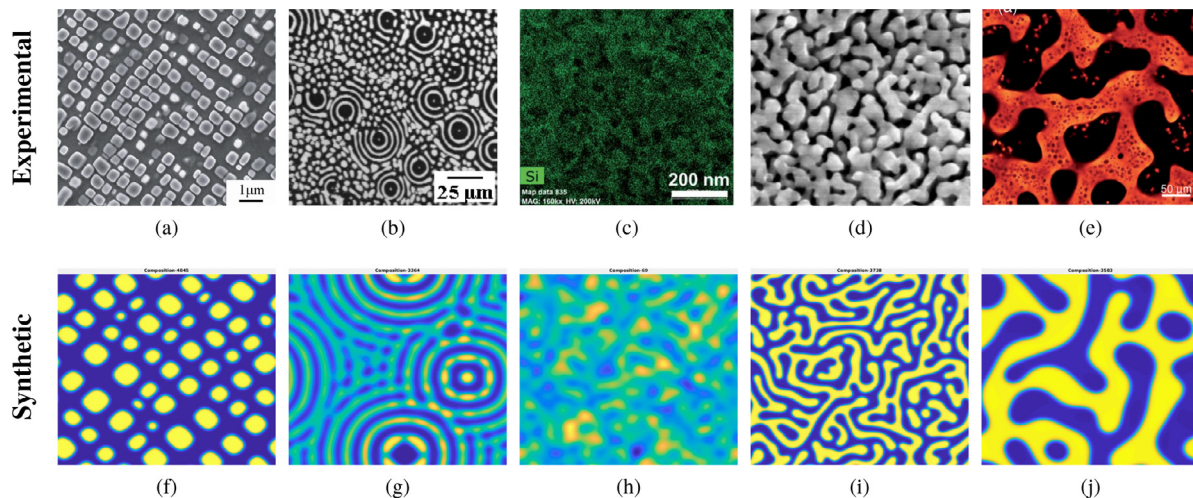


Fig. B.1. Selected experimental microstructures compared to the calculated ones. Experimental microstructures of a) SEM micrograph of $\gamma + \gamma'$ structure of Co-9.2Al-10.2W(at.%) reproduced from [91] with permission, b) ring-shaped pattern and droplets appeared in Polystyrene, reproduced from Ref. [92] with permission, c) $\text{Mg}_2\text{Si}_{0.3}\text{Sn}_{0.7}$ after high energy ball milling and isothermal heat treatment reproduced from Ref. [40] with permission, and d) SEM micrograph of the $\text{Cu}_{50}\text{Zr}_{45}\text{Al}_5$ bulk metallic glass immersed in the 0.05 M HF solution for 1 days, reproduced from Ref. [93] with permission. e) CLSM micrograph of phase-separated gel systems composed of whey protein isolate and gellan gum incubated at 5 °C, reproduced from Ref. [94] with permission. e,f,g,h) Selected synthetic microstructures corresponding to the experimental observations.

cuboidal γ' phase homogeneously precipitates in the γ (Al) matrix. This is also very similar to the morphology observed in many of the Ni-based superalloys. A similar synthetic microstructure where the precipitates are in the form of cuboids and oriented in a similar fashion is shown in Fig. B.1(f). Fig. B.1(b) illustrates a series of unusual dewetted patterns in polystyrene [92]. A series of ring-shaped patterns and droplets appeared in a specific region of this polystyrene sample annealed for 2 hours at 165 °C. Similarly, a microstructure where ring-shaped patterns appear during the process of separation is shown in Fig. B.1(g). Further investigation of the underlying reasons for formation of ring-shaped patterns will be provided in a feature study. Fig. B.1(c) shows an Electron Dispersive Spectroscopy (EDS) analysis of samples of $\text{Mg}_2\text{Si}_x\text{Sn}_{1-x}$ treated by high energy ball milling for 2 minutes with a subsequent 3 hours annealing at 720 °C [40]. The micrograph shows dull colors with smaller reddish and greenish regions suggesting a degree of mixing due to induced elastic strains and suppression of the miscibility gap. Likewise, the phase-field calculations shown in Fig. B.1(h) are obtained by the elastochemical modeling where the separation has been suppressed due to elastic effects. Fig. B.1(d) shows a microstructure of $\text{Cu}_{50}\text{Zr}_{45}\text{Al}_5$ bulk metallic glass immersed in the 0.05 M HF solution [93]. Fig. B.1(e) shows a gelation process in a biopolymer system. The biopolymer is a mixture of 3% w/w whey protein isolate and 0.04% w/w of gellan gum. Incubation at 5 °C resulted in bicontinuous morphology with interconnected microstructures of a polymer-rich phase and a serum phase (in dark color) [94]. Similar synthetic morphologies are represented in Fig. B.1(i) and Fig. B.1(j).

References

- [1] Y. Qian, C. Jackson, F. Giorgi, B. Booth, Q. Duan, C. Forest, D. Higdon, Z.J. Hou, G. Huerta, Uncertainty quantification in climate modeling and projection, *Bull. Am. Meteorol. Soc.* 97 (5) (2016) 821–824.
- [2] S. Gupta, D. Ghosh, Uncertainty Quantification in Structural Engineering: Current Status and Computational Challenges, in: *Uncertainty Quantification in Computational Science: Theory and Application in Fluids and Structural Mechanics*, World Scientific, 2017, pp. 119–149.
- [3] S. Ghoreishi, D. Allaire, Adaptive uncertainty propagation for coupled multidisciplinary systems, *AIAA J.* (2017) 3940–3950.
- [4] E. Begoli, T. Bhattacharya, D. Kusnezov, The need for uncertainty quantification in machine-assisted medical decision making, *Nature Mach. Intell.* 1 (1) (2019) 20.
- [5] J.H. Panchal, S.R. Kalidindi, D.L. McDowell, Key computational modeling issues in integrated computational materials engineering, *Comput.-Aided Des.* 45 (1) (2013) 4–25.
- [6] D.L. McDowell, S.R. Kalidindi, The materials innovation ecosystem: a key enabler for the materials genome initiative, *MRS Bull.* 41 (4) (2016) 326–337.
- [7] R. Arroyave, S. Shields, C.-N. Chang, D. Fowler, R. Malak, D. Allaire, Interdisciplinary research on designing engineering material systems: results from a national science foundation workshop, *J. Mech. Des.* (2018).
- [8] R. Arroyave, D.L. McDowell, Systems approaches to materials design: past, present, and future, *Annu. Rev. Mater. Res.* 49 (2019).
- [9] J. Allison, Integrated computational materials engineering: a perspective on progress and future steps, *JOM* 63 (4) (2011) 15–18.
- [10] X. Liu, D. Furrer, J. Kisters, J. Holmes, Vision 2040: roadmap for integrated, multiscale modeling and simulation of materials and systems, 2018, (????).
- [11] E. Weinan, *Principles of Multiscale Modeling*, Cambridge University Press, 2011.
- [12] M. Sanghvi, P. Honarmandi, V. Attari, T. Duong, R. Arroyave, D.L. Allaire, Uncertainty propagation via probability measure optimized importance weights with application to parametric materials models, in: *AIAA Scitech 2019 Forum*, 2019, p. 0967.
- [13] W.K. Liu, L. Siad, R. Tian, S. Lee, D. Lee, X. Yin, W. Chen, S. Chan, G.B. Olson, L.-E. Lindgen, et al., Complexity science of multiscale materials via stochastic computations, *Int. J. Numer. Methods Eng.* 80 (6–7) (2009) 932–978.
- [14] B. Kouchmeshky, N. Zabaras, Microstructure model reduction and uncertainty quantification in multiscale deformation processes, *Comput. Mater. Sci.* 48 (2) (2010) 213–227.
- [15] M. Koslowski, A. Strachan, Uncertainty propagation in a multiscale model of nanocrystalline plasticity, *Reliab. Eng. Syst. Saf.* 96 (9) (2011) 1161–1170.
- [16] S. Salehghaffari, M. Rais-Rohani, E. Marin, D. Bammann, A new approach for determination of material constants of internal state variable based plasticity models and their uncertainty quantification, *Comput. Mater. Sci.* 55 (2012) 237–244.
- [17] P. Honarmandi, A. Solomou, R. Arroyave, D.C. Lagoudas, Uncertainty quantification of the parameters and predictions of a phenomenological constitutive model for thermally induced phase transformation in ni-Ti shape memory alloys, *Modell. Simul. Mater. Sci. Eng.* (2019).
- [18] N. Saunders, A.P. Miodownik, *CALPHAD (calculation of phase diagrams): A Comprehensive Guide*, 1, Elsevier, 1998.
- [19] I. Steinbach, B. Böttger, J. Eiken, N. Warnken, S. Fries, Calphad and phase-field modeling: a successful liaison, *J. Phase Equilibria/Diffus.* 28 (1) (2007) 101–106.
- [20] M. Stan, B.J. Reardon, A Bayesian approach to evaluating the uncertainty of thermodynamic data and phase diagrams, *Calphad* 27 (3) (2003) 319–323.
- [21] P. Honarmandi, T.C. Duong, S.F. Ghoreishi, D. Allaire, R. Arroyave, Bayesian uncertainty quantification and information fusion in CALPHAD-based thermodynamic modeling, *Acta Mater.* 164 (2019) 636–647.
- [22] W.D. Thomison, D.L. Allaire, A model reification approach to fusing information from multifidelity information sources, in: *19th AIAA non-deterministic approaches conference*, 2017, p. 1949.
- [23] P. Honarmandi, N. Paulson, R. Arroyave, M. Stan, Uncertainty quantification and propagation in CALPHAD modelling, *Modell. Simul. Mater. Sci. Eng.* (2019).
- [24] R.A. Otis, Z.-K. Liu, High-throughput thermodynamic modeling and uncertainty quantification for ICME, *JOM* 69 (5) (2017) 886–892.

- [25] B. Bocklund, R. Otis, A. Egorov, A. Obaied, I. Roslyakova, Z.-K. Liu, ESPEI For efficient thermodynamic database development, modification, and uncertainty quantification: application to Cu–Mg, *MRS Commun.* (2019) 1–10.
- [26] S.-G. Kim, W.T. Kim, T. Suzuki, Phase-field model for binary alloys, *Phys. Rev. E* 60 (6) (1999) 7186.
- [27] I. Steinbach, L. Zhang, M. Plapp, Phase-field model with finite interface dissipation, *Acta Mater.* 60 (6–7) (2012) 2689–2701.
- [28] V. Attari, S. Ghosh, T. Duong, R. Arroyave, On the interfacial phase growth and vacancy evolution during accelerated electromigration in Cu/Sn/Cu microjoints, *Acta Mater.* 160 (2018) 185–198.
- [29] V. Attari, A. Cruzado, R. Arroyave, Exploration of the microstructure space in TiAlZrN ultra-hard nanostructured coatings, *Acta Mater.* 174 (2019) 459–476.
- [30] K. Karayagiz, L. Johnson, R. Seede, V. Attari, B. Zhang, X. Huang, S. Ghosh, T. Duong, I. Karaman, A. Elwany, et al., Finite interface dissipation phase field modeling of Ni–Nb under additive manufacturing conditions, *arXiv:1906.10200* (2019).
- [31] T. Duong, R.E. Hackenberg, V. Attari, A. Landa, P.E. Turchi, R. Arroyave, Probing discontinuous precipitation in U–Nb, *arXiv:1907.00918* (2019).
- [32] N. Wang, S. Rokkam, T. Hochrainer, M. Pernice, A. El-Azab, Asymptotic and uncertainty analyses of a phase field model for void formation under irradiation, *Comput. Mater. Sci.* 89 (2014) 165–175.
- [33] L.S. Leon, R.C. Smith, W.S. Oates, P. Miles, Identifiability and active subspace analysis for a polydomain ferroelectric phase field model, in: *ASME 2017 Conference on Smart Materials, Adaptive Structures and Intelligent Systems*, American Society of Mechanical Engineers, 2017.
- [34] B. Böttger, R. Altenfeld, G. Laschet, G. Schmitz, B. Stöhr, B. Burbaum, An ICME process chain for diffusion brazing of alloy 247, *Integr. Mater. Manuf. Innov.* 7 (2) (2018) 70–85.
- [35] S.H. Lee, W. Chen, A comparative study of uncertainty propagation methods for black-box-type problems, *Struct. Multidiscip. Optim.* 37 (3) (2009) 239.
- [36] P. Honarmandi, A. Solomou, R. Arroyave, D. Lagoudas, Parametric analysis of a phenomenological constitutive model for thermally induced phase transformation in ni–Ti shape memory alloys, *arXiv:1808.07377* (2018).
- [37] G. Tapia, L. Johnson, B. Franco, K. Karayagiz, J. Ma, R. Arroyave, I. Karaman, A. Elwany, Bayesian calibration and uncertainty quantification for a physics-based precipitation model of nickel–titanium shape-memory alloys, *J. Manuf. Sci. Eng.* 139 (7) (2017) 071002.
- [38] A. Badillo, C. Beckermann, Phase-field simulation of the columnar-to-equiaxed transition in alloy solidification, *Acta Mater.* 54 (8) (2006) 2015–2026.
- [39] S. Wang, N. Mingo, Improved thermoelectric properties of $\text{Mg}_2\text{Si}_x\text{Ge}_y\text{Sn}_{1-x-y}$ nanoparticle-in-alloy materials, *Appl. Phys. Lett.* 94 (20) (2009) 203109.
- [40] S. Yi, V. Attari, M. Jeong, J. Jian, S. Xue, H. Wang, R. Arroyave, C. Yu, Strain-induced suppression of the miscibility gap in nanostructured Mg_2Si – Mg_2Sn solid solutions, *J. Mater. Chem. A* 6 (36) (2018) 17559–17570.
- [41] T. Aizawa, R. Song, Mechanically induced reaction for solid-state synthesis of Mg_2Si and Mg_2Sn , *Intermetallics* 14 (4) (2006) 382–391.
- [42] Y. Noda, H. Kon, Y. Furukawa, N. Otsuka, I.A. Nishida, K. Masumoto, Preparation and thermoelectric properties of $\text{Mg}_2\text{Si}_{1-x}\text{Ge}_x$ ($x = 0.0 \sim 0.4$) solid solution semiconductors, *Mater. Trans., JIM* 33 (9) (1992) 845–850.
- [43] N.A. Heinz, T. Ikeda, Y. Pei, G.J. Snyder, Applying quantitative microstructure control in advanced functional composites, *Adv. Funct. Mater.* 24 (15) (2014) 2135–2153.
- [44] H. Balout, P. Boulet, M.-C. Record, Thermoelectric properties of Sn-containing Mg_2Si nanostructures, *J. Phys. Chem. C* 119 (31) (2015) 17515–17521.
- [45] S. Gorsse, S. Vives, P. Bellanger, D. Riou, L. Laversenne, S. Miraglia, D.R. Clarke, Multi-scale architected thermoelectric materials in the $\text{Mg}_2(\text{Si}, \text{Sn})$ system, *Mater. Lett.* 166 (2016) 140–144.
- [46] S. Gorsse, P. Bauer Pereira, R. Decourt, E. Sellier, Microstructure engineering design for thermoelectric materials: an approach to minimize thermal diffusivity, *Chem. Mater.* 22 (3) (2009) 988–993.
- [47] W. Wang, L. Yang, N. Wang, H. Zhang, Y. Jia, Design of a thermoelectric material using the calphad technique: Thermodynamic reassessment of the Al–Sb–Zn system, *J. Electron. Mater.* 47 (1) (2018) 261–271.
- [48] D.U. Furrer, Application of phase-field modeling to industrial materials and manufacturing processes, *Curr. Opin. Solid State Mater. Sci.* 15 (3) (2011) 134–140.
- [49] G.J. Schmitz, U. Pahl, Toward a virtual platform for materials processing, *JOM* 61 (5) (2009) 19–23.
- [50] Y.C. Yabansu, P. Steinmetz, J. Hötzer, S.R. Kalidindi, B. Nestler, Extraction of reduced-order process-structure linkages from phase-field simulations, *Acta Mater.* 124 (2017) 182–194.
- [51] D. Wang, R. Shi, Y. Zheng, R. Banerjee, H.L. Fraser, Y. Wang, Integrated computational materials engineering (icme) approach to design of novel microstructures for Ti-alloys, *JOM* 66 (7) (2014) 1287–1298.
- [52] G.J. Schmitz, B. Böttger, M. Apel, Microstructure modeling in ICME settings, in: *Proceedings of the 3rd World Congress on Integrated Computational Materials Engineering (ICME 2015)*, Springer, 2015, pp. 165–172.
- [53] G.J. Snyder, E.S. Toberer, Complex Thermoelectric Materials, in: *Materials For Sustainable Energy: A Collection of Peer-Reviewed Research and Review Articles from Nature Publishing Group, World Scientific*, 2011, pp. 101–110.
- [54] J. Yang, H.-L. Yip, A.K.-Y. Jen, Rational design of advanced thermoelectric materials, *Adv. Energy Mater.* 3 (5) (2013) 549–565.
- [55] G. Tan, L.-D. Zhao, M.G. Kanatzidis, Rationally designing high-performance bulk thermoelectric materials, *Chem. Rev.* 116 (19) (2016) 12123–12149.
- [56] P. Gorai, D. Gao, B. Ortiz, S. Miller, S.A. Barnett, T. Mason, Q. Lv, V. Stevanović, E.S. Toberer, TE Design lab: a virtual laboratory for thermoelectric material design, *Comput. Mater. Sci.* 112 (2016) 368–376.
- [57] A.S. Tazebay, S.-I. Yi, J.K. Lee, H. Kim, J.-H. Bahk, S.L. Kim, S.-D. Park, H.S. Lee, A. Shakouri, C. Yu, Thermal transport driven by extraneous nanoparticles and phase segregation in nanostructured $\text{Mg}_2(\text{Si}, \text{Sn})$ and estimation of optimum thermoelectric performance, *ACS Appl. Mater. Interf.* 8 (11) (2016) 7003–7012.
- [58] W. Liu, X. Tan, K. Yin, H. Liu, X. Tang, J. Shi, Q. Zhang, C. Uher, Convergence of conduction bands as a means of enhancing thermoelectric performance of n-type $\text{Mg}_2\text{Si}_{1-x}\text{Sn}_x$ solid solutions, *Phys. Rev. Lett.* 108 (16) (2012) 166601.
- [59] Q. Zhang, J. He, T. Zhu, S. Zhang, X. Zhao, T. Tritt, High figures of merit and natural nanostructures in $\text{Mg}_2\text{Si}_{0.4}\text{Sn}_{0.6}$ based thermoelectric materials, *Appl. Phys. Lett.* 93 (10) (2008) 102109.
- [60] V. Zaitsev, M. Fedorov, E. Gurieva, I. Eremin, P. Konstantinov, A.Y. Samunin, M. Vedernikov, Highly effective $\text{Mg}_2\text{Si}_{1-x}\text{Sn}_x$ thermoelectrics, *Phys. Rev. B* 74 (4) (2006) 045207.
- [61] G. Nolas, D. Wang, M. Beekman, Transport properties of polycrystalline $\text{Mg}_2\text{Si}_{1-y}\text{Sb}_y$ ($0 \leq y < 0.4$), *Phys. Rev. B* 76 (23) (2007) 235204.
- [62] S. Vives, P. Bellanger, S. Gorsse, C. Wei, Q. Zhang, J.-C. Zhao, Combinatorial approach based on interdiffusion experiments for the design of thermoelectrics: application to the $\text{mg}_2(\text{si}, \text{sn})$ alloys, *Chem. Mater.* 26 (15) (2014) 4334–4337.
- [63] A. Kozlov, J. Gröbner, R. Schmid-Fetzer, Phase formation in Mg–Sn–Si and Mg–Sn–Si–Ca alloys, *J. Alloys Compd.* 509 (7) (2011) 3326–3337.
- [64] R. Vienneis, C. Colinet, P. Jund, J.-C. Tedenac, Phase stability of ternary antiferroite type compounds in the quasi-binary systems Mg_2X – Mg_2Y (X, Y = Si, Ge, Sn) via ab-initio calculations, *Intermetallics* 31 (2012) 145–151.
- [65] E. Nikitin, V. Bazanov, V. Tarasov, Thermoelectric properties of Mg_2Si – Mg_2Sn solid solutions, *Sov. Phys. Solid State* 3 (1961) 2648–2652.
- [66] G. Polymeris, N. Vlachos, A. Khan, E. Hatzikrioniotis, C.B. Lioutas, A. Delimitis, E. Pavlidou, K. Paraskevopoulos, T. Kyratsi, Nanostructure and doping stimulated phase separation in high-ZT $\text{Mg}_2\text{Si}_{0.55}\text{Sn}_{0.4}\text{Ge}_{0.05}$ compounds, *Acta Mater.* 83 (2015) 285–293.
- [67] D. Rowe, V. Shukla, N. Savvides, Phonon scattering at grain boundaries in heavily doped fine-grained silicon–germanium alloys, *Nature* 290 (5809) (1981) 765.
- [68] D. Medlin, G. Snyder, Interfaces in bulk thermoelectric materials: a review for current opinion in colloid and interface science, *Curr. Opin. Colloid Interf. Sci.* 14 (4) (2009) 226–235.
- [69] T. Ikeda, L.A. Collins, V.A. Ravi, F.S. Gascoin, S.M. Haile, G.J. Snyder, Self-assembled nanometer lamellae of thermoelectric pbt and sb_2te_3 with epitaxy-like interfaces, *Chem. Mater.* 19 (4) (2007) 763–767.
- [70] Y. Pei, N.A. Heinz, A. LaLonde, G.J. Snyder, Combination of large nanostructures and complex band structure for high performance thermoelectric lead telluride, *Energy Environ. Sci.* 4 (9) (2011) 3640–3645.
- [71] P. Gorai, V. Stevanović, E.S. Toberer, Computationally guided discovery of thermoelectric materials, *Nat. Rev. Mater.* 2 (9) (2017) 17053.
- [72] L.Q. Chen, J. Shen, Applications of semi-implicit fourier-spectral method to phase field equations, *Comput. Phys. Commun.* 108 (2–3) (1998) 147–158.
- [73] A. Gelman, H.S. Stern, J.B. Carlin, D.B. Dunson, A. Vehtari, D.B. Rubin, *Bayesian Data Analysis*, Chapman and Hall/CRC, 2013.
- [74] D. Foreman-Mackey, D.W. Hogg, D. Lang, J. Goodman, emcee: the MCMC hammer, *Publ. Astron. Soc. Pac.* 125 (925) (2013) 306.
- [75] G.J. McRae, J.W. Tilden, J.H. Seinfeld, Global sensitivity analysis—a computational implementation of the Fourier amplitude sensitivity test (FAST), *Comput. Chem. Eng.* 6 (1) (1982) 15–25.
- [76] D. Huntington, C. Lyrintzis, Improvements to and limitations of latin hypercube sampling, *Probab. Eng. Mech.* 13 (4) (1998) 245–253.
- [77] H. Haario, M. Laine, A. Mira, E. Saksman, Dram: efficient adaptive MCMC, *Stat. Comput.* 16 (4) (2006) 339–354.
- [78] H. Haario, E. Saksman, J. Tamminen, et al., An adaptive metropolis algorithm, *Bernoulli* 7 (2) (2001) 223–242.
- [79] K. Choi, Y. Noh, I. Lee, Reliability-based design optimization with confidence level for problems with correlated input distributions, 6th China–Japan–Korea Joint Symposium on Optimization of Structural and Mechanical Systems, 2010.
- [80] D.J. Battle, P. Gerstoft, W.S. Hodgkiss, W. Kuperman, P.L. Nielsen, Bayesian model selection applied to self-noise geoaoustic inversion, *J. Acoust. Soc. Am.* 116 (4) (2004) 2043–2056.
- [81] E. Fried, M.E. Gurtin, Coherent solid-state phase transitions with atomic diffusion: a thermomechanical treatment, *J. Stat. Phys.* 95 (5–6) (1999) 1361–1427.
- [82] M.V. Karsanina, K.M. Gerke, E.B. Skvortsova, D. Mallants, Universal spatial correlation functions for describing and reconstructing soil microstructure, *PLoS ONE* 10 (5) (2015) e0126515.
- [83] D.T. Fullwood, S.R. Niezgod, S.R. Kalidindi, Microstructure reconstructions from 2-point statistics using phase-recovery algorithms, *Acta Mater.* 56 (5) (2008) 942–948.
- [84] J. Shotton, M. Johnson, R. Cipolla, Semantic texton forests for image categorization and segmentation, in: *2008 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2008, pp. 1–8.
- [85] G. Csúrk, C. Dance, L. Fan, J. Willamowski, C. Bray, Visual categorization with bags of keypoints, in: *Workshop on statistical learning in computer vision*, ECCV, I, Prague, 2004, pp. 1–2.
- [86] D.G. Lowe, Distinctive image features from scale-invariant keypoints, *Int. J. Comput. Vis.* 60 (2) (2004) 91–110.
- [87] Y. Ham, M. Kamari, Automated content-based filtering for enhanced vision-based documentation in construction toward exploiting big visual data from drones, *Autom. Constr.* 105 (2019) 102831.

- [88] K. Rajan, Informatics for Materials Science and Engineering: Data-driven Discovery for Accelerated Experimentation and Application, Butterworth-Heinemann, 2013.
- [89] C.K. Williams, C.E. Rasmussen, Gaussian Processes for Machine Learning, 2, MIT Press Cambridge, MA, 2006.
- [90] V. Attari, D.J. Saucedo, Open Phase-field Microstructure Database (OPMD), 2019. <http://microstructures.net>.
- [91] K. Ishida, Intermetallic compounds in Co-base alloys—phase stability and application to superalloys, MRS Online Proc. Library Arch. 1128 (2008).
- [92] J. Peng, R. Xing, Y. Wu, B. Li, Y. Han, W. Knoll, D.H. Kim, Dewetting of thin polystyrene films under confinement, *Langmuir* 23 (5) (2007) 2326–2329.
- [93] C. Wang, M. Li, M. Zhu, H. Wang, C. Qin, W. Zhao, Z. Wang, Controlling the mechanical properties of bulk metallic glasses by superficial dealloyed layer, *Nanomaterials* 7 (11) (2017) 352.
- [94] S. Wassén, R. Bordes, T. Gebäck, D. Bernin, E. Schuster, N. Lorén, A.-M. Hermansson, Probe diffusion in phase-separated bicontinuous biopolymer gels, *Soft Matter* 10 (41) (2014) 8276–8287.