

PAPER

## Uncertainty analysis of microsegregation during laser powder bed fusion

To cite this article: Supriyo Ghosh *et al* 2019 *Modelling Simul. Mater. Sci. Eng.* **27** 034002

View the [article online](#) for updates and enhancements.

### Recent citations

- [Statistical modelling of microsegregation in laser powder-bed fusion](#)  
Supriyo Ghosh *et al*
- [Microsegregation behavior of Inconel 718 superalloy prepared by electron beam smelting layered solidification technology](#)  
Longhai Zhao *et al*
- [Review: Materials Ecosystem for Additive Manufacturing Powder Bed Fusion Processes](#)  
Behrang Poorganji *et al*



**IOP | ebooks™**

Bringing together innovative digital publishing with leading authors from the global scientific community.

Start exploring the collection—download the first chapter of every title for free.

# Uncertainty analysis of microsegregation during laser powder bed fusion

Supriyo Ghosh<sup>1</sup> , Mohamad Mahmoudi<sup>2</sup>, Luke Johnson<sup>1</sup>, Alaa Elwany<sup>1,2</sup>, Raymundo Arroyave<sup>1,2,3</sup> and Douglas Allaire<sup>3</sup>

<sup>1</sup> Department of Materials Science & Engineering, Texas A&M University, College Station, TX 77843, United States of America

<sup>2</sup> Department of Industrial & Systems Engineering, Texas A&M University, College Station, TX 77843, United States of America

<sup>3</sup> Department of Mechanical Engineering, Texas A&M University, College Station, TX 77843, United States of America

E-mail: [gsupriyo2004@gmail.com](mailto:gsupriyo2004@gmail.com) and [supriyo.ghosh@tamu.edu](mailto:supriyo.ghosh@tamu.edu)

Received 1 October 2018, revised 9 January 2019

Accepted for publication 24 January 2019

Published 25 February 2019



CrossMark

## Abstract

Quality control in additive manufacturing can be achieved through variation control of the quantity of interest (QoI). We choose in this work the microstructural microsegregation to be our QoI. Microsegregation results from the spatial redistribution of a solute element across the solid–liquid interface that forms during solidification of an alloy melt pool during the laser powder bed fusion process. Since the process as well as the alloy parameters contribute to the statistical variation in microstructural features, uncertainty analysis of the QoI is essential. High-throughput phase-field simulations estimate the solid–liquid interfaces that grow for the melt pool solidification conditions that were estimated from finite element simulations. Microsegregation was determined from the simulated interfaces for different process and alloy parameters. Correlation, regression, and surrogate model analyses were used to quantify the contribution of different sources of uncertainty to the QoI variability. We found negligible contributions of thermal gradient and Gibbs–Thomson coefficient and considerable contributions of solidification velocity, liquid diffusivity, and segregation coefficient on the QoI. Cumulative distribution functions and probability density functions were used to analyze the distribution of the QoI during solidification. Our approach, for the first time, identifies the uncertainty sources and frequency densities of the QoI in the solidification regime relevant to additive manufacturing.

Keywords: additive manufacturing, phase-field modeling, surrogate modeling, microsegregation, solidification

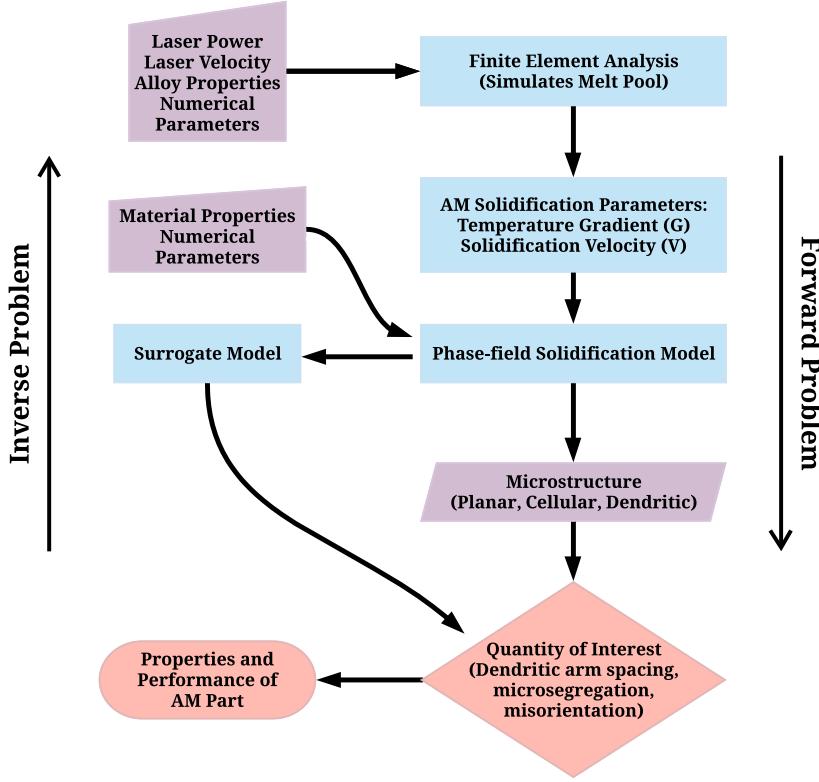
(Some figures may appear in colour only in the online journal)

## 1. Introduction

Additive manufacturing (AM) processes are increasingly becoming pervasive due to their ability to produce intricate parts with improved properties compared to traditional manufacturing processes and are therefore widely being used in aerospace and automotive industries [1, 2]. The AM process studied in the present work is laser powder bed fusion (LPBF). The material chosen for research is a Ni–Nb alloy, which is a binary analog of Inconel 718, that finds application in jet-engine and gas-turbine components [1, 2]. During the LPBF process, the laser beam moves across the alloy powder with a fixed velocity, resulting in a liquid melt pool that solidifies into a complex solid–liquid interface. The morphology of the solid–liquid interface is determined by the solidification conditions in the melt pool, namely, temperature gradient  $G$  and solidification velocity  $V$  [3]. The typical solidification morphologies are planar, cellular, and dendritic in which the key microstructural features such as spacing, segregation, and orientation determine the properties of an AM component.

The primary barrier to the widespread adoption of the new classes of AM materials is the lack of confidence in the product quality [4–6], which is due to the variability present at various stages of the manufacturing process. Development of a quality control approach based on uncertainty quantification (UQ) of the process through multiscale simulations is a key to resolve this issue. The flowchart in figure 1 summarizes our multiscale uncertainty analysis approach. A macro-scale finite element analysis (FEA) model simulates the LPBF process where the potential sources of uncertainty are the laser processing parameters, material properties, and boundary conditions. The melt pool profile obtained from FEA simulations is used to determine the solidification parameters as they are inputs to the microstructure model. A phase-field model [7–9] is used for the simulation of microstructure, from which the quantity of interests (QoIs) are determined. Phase-field models can be computationally intensive as the simulations [10–12] often run from hours to days to produce informative microstructures. Surrogate models are inexpensive approximations to the original computer model ([13] and the references within). Therefore, a suitable surrogate model could potentially substitute for the expensive phase-field simulations.

The microstructural features that develop during LPBF solidification process are statistically variable in several aspects [14, 15]. The solidification morphology and the distribution of the grain size can vary from region to region within the microstructure. The spatial distribution of the solute composition in the as-built alloy samples can be severely inhomogeneous due to rapid cooling rates ( $\dot{T} = G \times V$ ). The misorientation of the solidifying grains with respect to the build direction is frequently observed in multi-layer multi-pass solidification microstructures [16]. The above microstructural features further vary as a function of the alloy properties, such as the segregation coefficient, the diffusivity of the liquid, and the Gibbs–Thomson coefficient. All the above issues lead to inhomogeneous mechanical properties of the solid material. From a modeling perspective, numerical parameters such as interface width and mesh size contribute to uncertainties. It is entirely unavoidable to eliminate the microstructural variabilities due to the non-equilibrium nature of the solidification during LPBF [4, 6]. The incorporation of non-equilibrium physics adds further



**Figure 1.** Summary flowchart of uncertainty propagation in a multi-level simulation framework.

computational cost to the original computer simulation models, thus requiring efficient surrogate models to approximate those.

We chose to study the microsegregation (i.e. the spatial distribution of the solute across the interface) in LPBF solidification microstructures as our QoI. Microsegregation severely affects the yield, tensile, and fatigue strengths of the material and can be estimated from a microstructure in a relatively small domain. Our strategy in this work is to use a microstructure model to determine the QoI statistics in LPBF microstructures and to apply the knowledge of UQ to understand the contribution of various sources to the QoI variability. Microstructure evolution during a low-velocity solidification process possesses much less variability of the microstructural features compared to that of the same during the LPBF process and hence is often ignored in the casting solidification literature. Conversely, uncertainty analysis is essential for the quantification of LPBF microstructures due to significant variability in its key features. It is, however, rarely addressed in the literature (for recent overviews, refer to [17–19]).

UQ frameworks of a FEA model for melt pool simulation and a cellular automata model for microstructure simulation were coupled very recently [17, 20]. The phase-field method is more efficient compared to the cellular automata method in that the phase-field method does not require explicit tracking of the solid–liquid interface and the difficulties associated with the estimation of interfacial curvature are handled efficiently [21, 22]. There is no study, as far as we are aware, that addresses uncertainty during the high-fidelity microstructure evolution

using the phase-field method. Further, one must be cautious in using statistics derived from limited input and output data during phase-field simulation of the LPBF process, as small datasets may not necessarily represent the overall processing as well as the solidification map. Therefore, the distribution derived from a large dataset could potentially be used as a metric to gain a deeper understanding of the physics and trends involved in the solidification process. Uncertainty analysis for LPBF is a relatively new research field with tremendous growth opportunities in the future. In this first approach, key questions we address in the microsegregation modeling in the solidification regime relevant to LPBF are: Which model parameter uncertainties affect microsegregation? How can statistical distributions be used to represent microsegregation? How do sample size and distribution affect microsegregation? Are surrogate models good enough to approximate phase-field models?

The remainder of the article is as follows. Simulation method, parameters, and microstructure analysis techniques are outlined in section 2. The sensitivity, surrogate, and frequency analyses of our phase-field simulation results are illustrated in section 3. A detailed discussion of our simulation results is given in section 4. Finally, a summary and outlook of the current work are given in section 5.

## 2. Simulation details

### 2.1. Method

Inconel 718 (IN718) is a multicomponent alloy with more than ten alloying elements making the microsegregation in the solidifying melt pool complex. Our strategy here is to conduct an uncertainty analysis of a binary approximation of the IN718 alloy, the understanding of which can be effectively extended for the modeling of a multicomponent analog using a multi-phase-field model that handles multiple components and phases [23–25]. Ni–Nb is the most important binary analog of IN718, since Nb segregates most severely from the solid to the liquid during directional solidification of the IN718 melt pool due to the smallest partition coefficient of Nb among all elements in IN718 and thus has a major role in determining the key solidification microstructural features as well as the subsequent solid–solid precipitation reactions.

We have used a phase-field solidification model detailed in [26, 27] to simulate the solid–liquid interfaces that evolve during directional solidification of a Ni-5 wt% Nb alloy. The model was developed to obtain a quantitative assessment of the microstructure information that results during directional growth of a binary alloy. The model is employed and validated extensively in the literature [9, 26, 28] and the simulation results agree excellently with the experimental measurements on metallic alloys in the low-velocity limit. Therefore, as a reference, such an established model is programmed to simulate the evolution of the non-conserved phase-field  $\phi$  and the conserved composition field  $c$  during the time( $t$ )-dependent solidification process in the melt pool. The phase-field  $\phi$  is a scalar-valued order parameter field which distinguishes the microstructure phases;  $\phi = 1$  in the solid,  $\phi = -1$  in the liquid and the solid–liquid interface is described by  $-1 < \phi < 1$ . This approach avoids explicit tracking of the interface and thus the complex solid–liquid interfaces are extracted in an efficient way. The effects of melt convection are not included, the diffusion of solute in the solid is neglected, the diffusion of heat is ignored, and local equilibrium at the solid–liquid interface is imposed in this model. The evolution equation for  $\phi$  can be written as:

$$\begin{aligned} \tau_0 a(\hat{q})^2 \frac{\partial \phi}{\partial t} = & W_0^2 \nabla \cdot [a(\hat{q})^2 \nabla \phi] + \sum_{i=1}^d \partial_i \left[ a(\hat{q}) \frac{\partial a(\hat{q})}{\partial (\partial_i \phi)} |\nabla \phi|^2 \right] \\ & + \phi - \phi^3 - \frac{\lambda}{1-k} (1-\phi^2)^2 \left[ \exp(u) - 1 + \frac{G(z-Vt)}{m_l c_0 / k} \right]. \end{aligned} \quad (1)$$

The dimensionless surface energy function  $a(\hat{q}) = 1 - \epsilon [3 - 4 \sum_{i=1}^d q_i^4]$  represents the  $d$ -dimensional fourfold anisotropy at the solid–liquid interface with a strength  $\epsilon$  and  $q_i$  is the interface normal vector pointing into liquid along the Cartesian direction  $i$  in the lab frame of reference. Alloy composition  $c_0$ , liquidus slope  $m_l$ , and equilibrium partition coefficient (or, segregation coefficient)  $k$  can be approximated from a Ni–Nb phase diagram [29, 30]. The dimensionless chemical potential  $u$  is given by  $\ln \left( \frac{2ck/c_0}{1+k-(1-k)\phi} \right)$ . A frozen-temperature approximation [26] is applied in which the temperature gradient  $G$  is translated along the  $z$  (growth) axis with a velocity  $V$ .

The evolution equation for  $c$  can be written as:

$$\frac{\partial c}{\partial t} = -\nabla \cdot \left[ -\frac{1}{2} (1+\phi) D c \exp(u)^{-1} \nabla \exp(u) + a_t W_0 (1-k) \exp(u) \frac{\partial \phi}{\partial t} \frac{\nabla \phi}{|\nabla \phi|} \right], \quad (2)$$

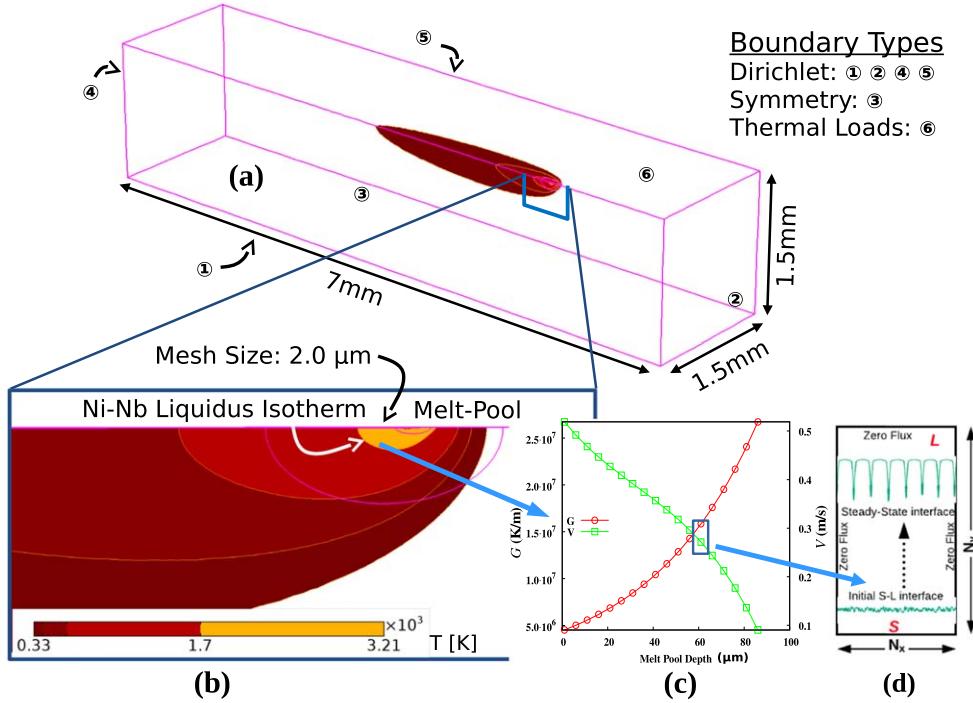
where the first term inside the square bracket represents a standard Fickian diffusion flux, and the second term is the anti-trapping solute flux term that prevents any artificial solute trapping to occur at the simulated solid–liquid interface and thus the solute redistribution across the interface becomes efficient. The value of  $a_t = 1/(2\sqrt{2})$  is based on the thin interface model by Karma [27].  $D$  is the diffusivity of solute in the liquid.

The numerical parameters in the model, the interface thickness  $W_0$ , the phase-field relaxation time  $\tau_0$ , and the dimensionless coupling constant  $\lambda$ , are linked to material parameters via the chemical capillary length  $d_0 = 0.8839 W_0 / \lambda$  and the timescale for diffusion  $\tau_0 = 0.6267 \lambda W_0^2 / D$  that use a thin-interface analysis in order to make the interface kinetics vanish. Both  $W_0$  and  $\tau_0$  values are used to render all the simulation parameters dimensionless.

## 2.2. Parameters and procedures

**2.2.1. Finite element analysis.** The solidification parameters in equation (1) were estimated using 3D FEA simulations carried out within the COMSOL Multiphysics [31] heat transfer module. A single-track laser scan in the length direction of a rectangular parallelepiped specimen of Ni–Nb substrate at an initial uniform temperature of 298 K was modeled (figure 2). The laser beam power distribution was assumed to be Gaussian with a beam diameter of 80  $\mu\text{m}$ . The power absorption coefficient was considered to be phase-dependent following [32]. The absorptivity values of solid, liquid, and vapor phases was 0.3, 0.3, and 0.6, respectively. The governing equations for the conservation of energy during the laser heat distribution within the material can be found in [33]. The boundary conditions listed as thermal loads (figure 2(a)) included heat transfer considerations of the deposited laser beam, natural convection, radiation, and vaporization. Natural convection and radiation contributions were included through the standard implementations of each phenomenon typical of heat transfer analysis [33]. Vaporization was included through a loss of energy via the mass flux of vapor leaving the system which was calculated using the weld pool evaporative flux model developed in [34].

Heat transfer within the simulation domain was governed by conduction only, but included phase-transformation effects in the form of latent heat contributions to the specific



**Figure 2.** (a) Three-dimensional FEA simulations determine the temperature distribution in the resulting melt pool during LPBF (laser power of 200 W and beam speed of  $500 \text{ mm s}^{-1}$ ) of a Ni-5% Nb alloy. A typical melt pool profile is shown within the geometry of the simulation domain. (b) A 2D section of the 3D melt pool cut along its center in the length direction is presented to illustrate the temperature distribution. (c)  $G$  and  $V$  vary as a function of the melt pool depth. (d) These  $G$  and  $V$  values are the inputs to the phase-field simulation. Note that this schematic is one instance of multiple such parallel FEA and phase-field calculations that are being used for the uncertainty analysis of the LPBF microstructures.

heat during melting and vaporization, as well as phase-dependent material properties [33]. Thermophysical properties such as the bulk material density, latent heat, and specific heat of the solid, liquid, and vapor phases of the Ni–Nb alloy were calculated using an atomic composition based rule of mixtures of pure Ni and Nb properties. For a small alloying addition of Nb (5 wt%), the properties are essentially that of pure Ni which can be found in the Material Property Handbook [35, 36]. Liquid conductivity was calculated as twice that of the solid to approximate the effect of convection within the melt pool, as is common practice in heat transfer analysis [37]. Vapor phase conductivity was artificially enhanced to simulate the transmission of laser deposited energy through the phase into the substrate during vaporization. This enhanced vapor conductivity is unique to this model, but has similarities to the element birth and death method [38]. However, our approach is more physical in that it retains the vapor that forms in the region between the laser beam and the substrate material.

The LPBF process was simulated for a wide range of laser parameters that were based on our knowledge and machine specifications; laser power typically varied between 30 W and 300 W and beam speed typically varied between 0.1 and  $2.5 \text{ m s}^{-1}$ . A typical resultant thermal profile after the FEA simulation is shown in figure 2(b). The melt pool profile obtained from this simulation was used to extract the solidification parameters  $G$  and  $V$

**Table 1.** Assumed ranges for the phase-field model input parameters.

Input ( $X$ )	Minimum	Maximum
$G$ (K m $^{-1}$ )	$10^6$	$4 \times 10^7$
$V$ (m s $^{-1}$ )	0.01	2.5
$D$ (m $^2$ s $^{-1}$ )	$10^{-9}$	$9 \times 10^{-9}$
$\Gamma$ (K m)	$10^{-7}$	$4 \times 10^{-7}$
$k$ (dimensionless)	0.48	0.74

(following their mathematical expressions given in [3, 39]) at the liquidus temperature isotherm where the solid–liquid phase transformation begins. The ranges of  $G$  and  $V$  (figure 2(c)) estimated for various values of laser parameters approximate the entire LPBF solidification map (table 1). These  $G$  and  $V$  values are inputs to the phase-field model (figure 2(d)).

It should be noted that we ignore the effects of uncertainty regarding the material properties (e.g. laser absorptivity, thermal conductivity of the liquid), melt pool physics (e.g. computational fluid dynamics and Marangoni convection), and numerical parameters (e.g. mesh size) that affect the temperature distribution as well as the melt pool size and geometry during our FEA simulations; a detailed uncertainty analysis of which can be found in [40]. An arbitrary single set of laser parameters may not necessarily lead to ranges of  $G$  and  $V$  that approximate the overall LPBF process and solidification maps, on average. This is why we choose to use wide ranges of laser parameters in our FEA simulations to use the resultant ranges of  $G$  and  $V$  values for the subsequent phase-field simulations. A large number of input samples designed from the resultant ranges of  $G$  and  $V$  is used to perform uncertainty analysis efficiently in the present work.

**2.2.2. Phase-field.** Equations (1) and (2) are solved on a two-dimensional uniform mesh using a finite volume method and no-flux boundary conditions in all directions. The numerical values of the parameters used in the simulations are: interface width of 0.5 nm; grid spacing of 0.3 nm; and two different mesh sizes (in grid units) of  $600(N_x) \times 3000(N_y)$  and  $10(N_x) \times 3000(N_y)$  are used to simulate cellular and/or planar interfaces. The interface thickness and grid spacing values used in our simulations are small enough so that the results become independent of their values. Each simulation begins with a thin layer of solid at the bottom of the simulation box with an initial Nb composition of  $kc_0$  in the solid and  $c_0$  in the liquid. Small, random amplitude perturbations are applied at the initial solid–liquid interface, from which perturbations either grow with time and break into a cellular interface or decay with time resulting into a planar interface. The spatial redistribution of the resulting composition field  $c$  across the interface is referred to as microsegregation.

Three input material parameters, namely, the diffusivity of the liquid  $D$ , the Gibbs–Thomson coefficient  $\Gamma$ , and the segregation coefficient  $k$  are found to affect significantly on the variability of the phase-field model output. Gibbs–Thomson coefficient  $\Gamma$  is given by  $d_0 \Delta T_0$ , where  $d_0$  is the capillary length and  $\Delta T_0$  is the equilibrium freezing range of the alloy. Although not shown here, we found that the variabilities of  $c_0$ ,  $m_l$ , and  $\epsilon$  (refer to equation (1)) are not significant and hence are ignored following preliminary phase-field experiments. The specific ranges of the uncertain material properties are decided based on the literature review [29, 30, 35, 36, 41].

We use a Latin hypercube sampling (LHS) method [42] to sample  $N = 100$  process and alloy parameters evenly across all possible values, the statistical measures of which are listed in table 1. LHS partitions each input distribution into  $N$  intervals of equal probability and select one sample randomly from each interval to construct the input design space. This is more efficient compared to the Monte Carlo sampling technique that randomly selects a sample from an input distribution, leading to some intervals in the sample space with a clustered data and some intervals with no samples.

### 2.3. Analysis of uncertainty

We use a scatterplot to visualize the variability of  $n$  input process and material parameters  $\mathbf{X} = \{x_{i1}, \dots, x_{in}\}_{i=1}^N$  on the QoI  $\mathbf{Y}$ . We perform sensitivity analysis by determining the Pearson product-moment correlation coefficient  $\rho_{XY}$  for each  $\mathbf{X}$  and  $\mathbf{Y}$  combinations [43, 44]:

$$\text{Corr}(\mathbf{X}, \mathbf{Y}) = \rho_{XY} = \frac{\text{Covariance}(\mathbf{X}, \mathbf{Y})}{\sigma_X \sigma_Y}, \quad (3)$$

where  $\sigma_X$  and  $\sigma_Y$  are the standard deviations of  $\mathbf{X}$  and  $\mathbf{Y}$ , respectively. The value of  $\rho_{XY}$  ranges between  $-1$  and  $1$ ;  $\rho_{XY} = 0$  signifies no correlation between  $\mathbf{X}$  and  $\mathbf{Y}$ , and the absolute value of  $\rho_{XY}$  defines the relationship strength between the two. A value of  $\rho_{XY}$  close to  $0$  signifies that the variability of the input  $\mathbf{X}$  on  $\mathbf{Y}$  can be discarded. From such an analysis, the number of uncertain input parameters can be reduced and the parameters can be ranked according to their correlation strengths with  $\mathbf{Y}$ .

A regression analysis [43, 44] was used to obtain the line of best fit from a typical scatterplot representing  $\mathbf{X}$  versus  $\mathbf{Y}$ . Regression line equations, as a first approximation, would help to find a rough estimate of the QoI  $\mathbf{Y}$  for arbitrary values of the input choices  $\mathbf{X}$ . The calculations of mean ( $\mu$ ) and standard deviation ( $\sigma$ ) of  $\mathbf{X}$  and  $\mathbf{Y}$  are necessary to perform regression analysis, which can be expressed as:  $\mathbf{Y} = a\mathbf{X} + b$ , where  $a = \rho_{XY} \sigma_Y / \sigma_X$  and  $b = \mu_Y - a \mu_X$ .  $\sigma_X = \sqrt{\sum_i^N (x_i - \mu_X)^2 / N}$  and  $\sigma_Y = \sqrt{\sum_i^N (y_i - \mu_Y)^2 / N}$ , where  $x_i$  and  $y_i$  are the elements in  $\mathbf{X}$  and  $\mathbf{Y}$ , respectively.

A Gaussian process (GP) surrogate model [45] was used to approximate  $\mathbf{Y}$ . The surrogate model was built on the training dataset  $(\mathbf{X}, \mathbf{Y})$  provided by the phase-field simulations. The performance of the surrogate model was estimated by measuring the difference between phase-field model ( $y_i$ ) and surrogate model ( $\hat{y}_i$ ) predictions, referred to as the mean absolute predictive error (MAPE):

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad \forall i \in \{1, \dots, N\}. \quad (4)$$

Here  $y_i$  is  $i$ th element of the phase-field model output at an input  $\mathbf{x}_i$ , and  $\hat{y}_i$  is the  $i$ th element of the surrogate model prediction evaluated at the same input  $\mathbf{x}_i$ . Similarly, mean absolute percent error is given by  $\frac{1}{N} \sum_{i=1}^N \frac{|y_i - \hat{y}_i|}{y_i} \times 100$ .

Frequency analysis of  $\mathbf{Y}$  is performed using cumulative distribution (CDF) and probability density functions (PDFs) [43, 44]. Phase-field simulated  $\mathbf{Y}$  values are ranked  $M = 1$  through  $M = N = 100$  in the ascending order and then, for each  $\mathbf{Y}$  entry, the empirical CDF is calculated using the formula:  $M/N$  and plotted against  $\mathbf{Y}$ . The derivative at any point in this CDF plot is regarded as the probability of the associated PDF. The empirical CDF is fitted against different target CDFs such as normal, lognormal, gamma, beta, and Weibull in a MATLAB distribution fitting tool ‘dfittool’ [46] that uses maximum likelihood estimator to approximate  $\mathbf{Y}$ . This analysis helps to gain insights regarding the type of frequency

distribution the simulated  $\mathbf{Y}$  closely follows. Such an analysis was performed for various random samples to test whether there is any effect of sampling on the  $\mathbf{Y}$ .

### 3. Results

#### 3.1. General features

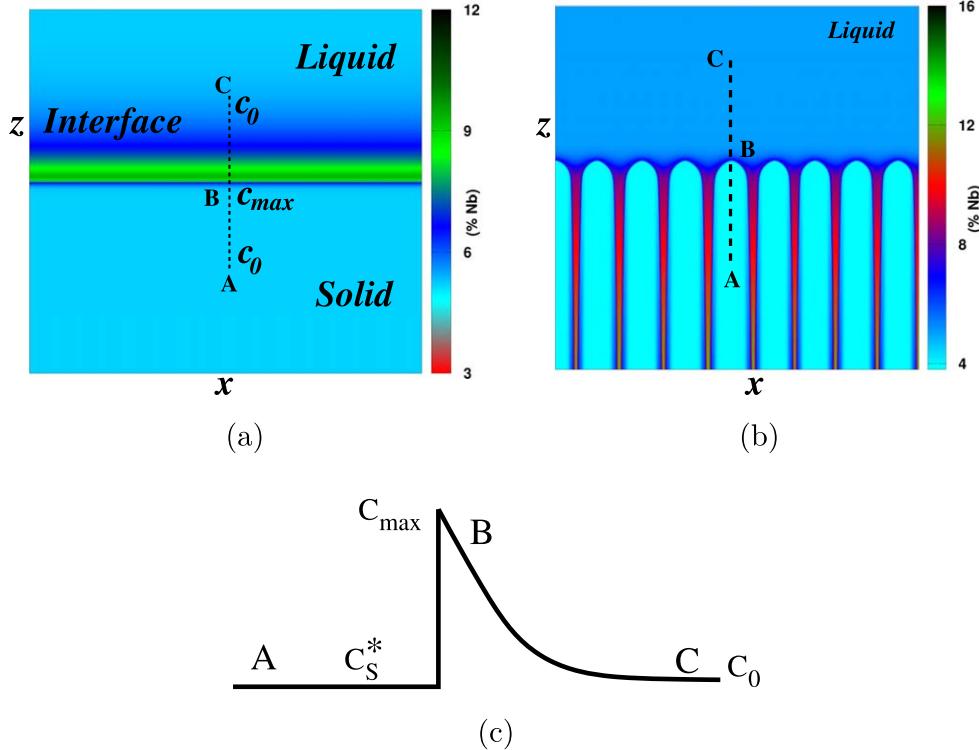
The solid–liquid interfaces in a solidifying melt pool grow in the liquid in the direction of the temperature gradient  $G$  at a rate  $V$ . The morphology of the solidifying interface can be estimated by two critical velocities [47], the constitutional supercooling velocity  $V_{cs} = GD/\Delta T_0$  and the absolute velocity  $V_{ab} = \Delta T_0 D/(k\Gamma)$ . When  $V$  is below  $V_{cs}$ , the interface morphology would be planar, for  $V_{cs} < V < V_{ab}$ , the interface would be cellular/dendritic, and, for  $V > V_{ab}$ , the interface becomes planar again.

The essence of the planar front and cellular solidification from the phase-field simulation is as follows. Simulation starts with the procedure described in section 2.2. The initial perturbation in the solid–liquid interface either decays or grows (referred to as the Mullins–Sekerka instability [48]) depending on the size of the domain and the strength of the perturbation, leading to either planar or cellular interface. Cellular solidification undergoes intermediate transient stages of growth by merging or splitting of the neighboring cells, which finally develop into a steady state cellular microstructure. At steady state, the solidification front grows with the temperature field at a constant velocity of  $V$ , which equals that estimated from the FEA simulation (refer section 2.2). Typical simulations of planar and cellular interfaces are shown in figure 3.

Let us consider that  $c_s^*$  is the composition of Nb in the solid and  $c_{\max}$  is the maximum composition of Nb in the liquid side of the interface. The scaled microsegregation is then determined by the solid composition ( $c_s^*$ ) to the liquid composition ( $c_{\max}$ ) at the interface, given by  $k_v = c_s^*/c_{\max}$ . For a planar interface,  $c_s^* = c_0$  (nominal composition). For a cellular interface,  $c_s^* < c_0$  for a Ni–Nb alloy. Note that there is an uncertainty associated with the determination of  $c_{\max}$ , this being a steeply peaked function, which we ignore in the present work. Since the diffusivity of the solute in the solid is several orders of magnitude smaller than that of the same in the liquid, we ignore diffusion in the solid and thus  $c_s^*$  remains constant in the solid. In our simulations, the solid–liquid interface is not in local equilibrium, since the  $k_v$  value extracted at the interface deviates from the equilibrium segregation coefficient  $k$ . In this context, an analytical approximation such as the Gulliver–Scheil equation [47] can predict the solute redistribution during planar solidification for the conditions of zero diffusion in the solid, infinite fast diffusion in the liquid, and local interface equilibrium. However, such an analytical approach would be even more strongly idealized compared to the present scenario where complex interface shapes form during non-equilibrium solidification and deviations from the phase-diagram are expected for arbitrary solidification conditions.

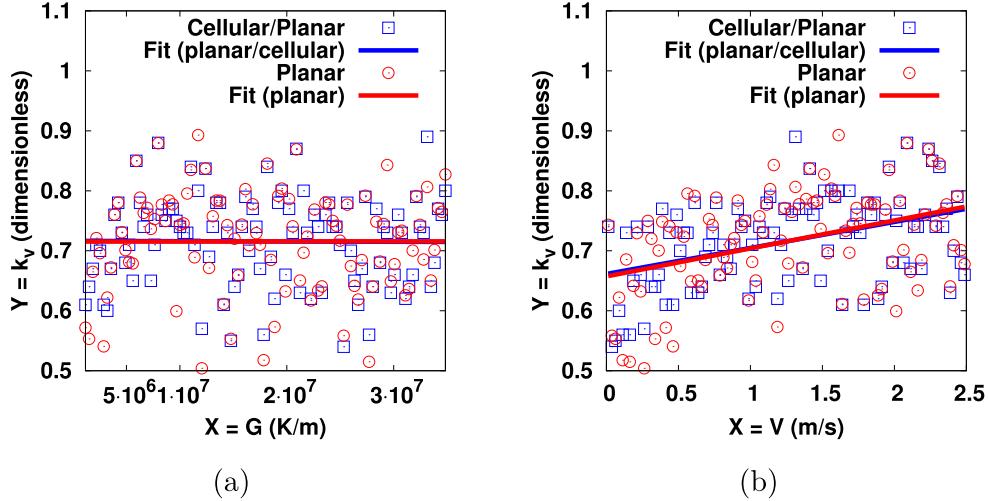
#### 3.2. Size effects

To quantify the effects of the domain size on the simulated microstructures, we perform phase-field simulations for two lateral sizes ( $N_x$ ) of the domain (refer to section 2.2). The small domain (ten grid points) is used to obtain a planar solid–liquid interface in all simulations. Simulations with the large lateral domain (600 grid points) resulted into either planar or cellular interface depending on the combinations of  $G$  and  $V$  inputs. The simulated planar and cellular interfaces are presented in figure 3. The values of  $k_v$  are extracted from these interfaces and are plotted against the model inputs  $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_5\} = \{G, V, D, \Gamma, k\}$  in

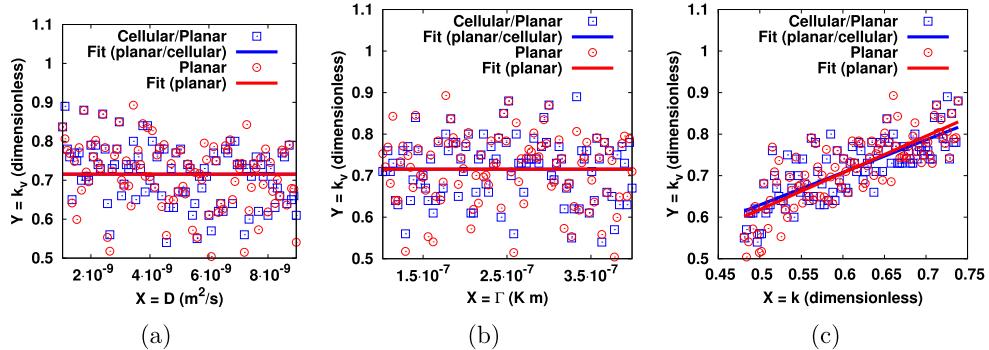


**Figure 3.** (a) Planar solid–liquid interface (not to scale) forms with a composition distribution that is shown using a color map. The corresponding Nb composition in both solid and liquid equals the nominal composition of the alloy  $c_0$ , and Nb enrichment at the interface is denoted by  $c_{max}$ . (b) Cellular interface that forms for a fixed value of ( $G$ ,  $V$ ) is illustrated using a color map. The characteristic compositions from A ( $c_s^*$ ), B ( $c_{max}$ ), and C ( $c_0$ ) are determined. (c) Composition distribution across a plane front or a cell tip determines the scaled microsegregation number  $k_v = c_s^*/c_{max}$ . For a planar interface,  $c_s^* = c_0$ .

figures 4 and 5. It is clear that the  $k_v$  extracted from planar and cellular solidification fronts are different when a single numerical experiment is considered that is performed with the same process and material parameters, but with different domain sizes. However, when the  $k_v$  data is considered from the entire population of phase-field experiments, the  $k_v$  distribution for both planar and cellular interfaces appears similar, on average, and the ‘average’ lines that represent these two  $k_v$  data set almost overlap with each other. Thus, the model output becomes size independent when the effects of the total population are considered. This is quantified using a sensitivity analysis in the following section. Since the distribution of  $k_v$  population derived from both interfaces is similar, at least for the present ranges of the solidification conditions, henceforth, we present the output data from planar solidification only. For the sake of completeness, we present the effect of mesh size and interface width on the simulated  $k_v$  values in figure 6. We find that below a mesh size of 0.5 nm and interface thickness of 0.5 nm simulation results become grid independent and therefore these numerical values are used for the present study.



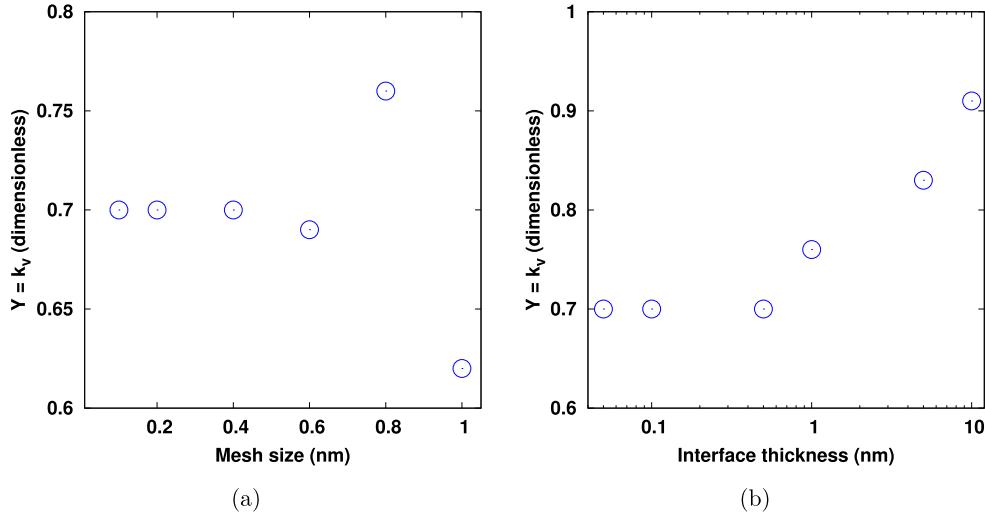
**Figure 4.** Scatterplots for (a)  $G$  and  $k_v$ , (b)  $V$  and  $k_v$  for planar and cellular interfaces are presented. These phase-field data are extracted from two different domain sizes. Simulations with the small domain resulted into a planar interface, and the simulations with the larger domain resulted into either cellular or planar interface. The lines of best fit estimated from the simulation outputs from two independent runs overlap, signifying the same output distribution of microsegregation, on average.



**Figure 5.** Scatterplots for (a)  $D_l$  and  $k_v$ , (b)  $\Gamma_{sl}$  and  $k_v$ , (c)  $k$  and  $k_v$  for planar and cellular interfaces are presented. These phase-field data are extracted from two different domain sizes. Simulations with the small domain resulted into a planar interface, and the simulations with the larger domain resulted into either cellular or planar interface. The lines of best fit estimated from the simulation outputs from two independent runs overlap, signifying the same output distribution of microsegregation, on average.

### 3.3. Sensitivity analysis

For a sensitivity analysis, the scatterplots representing  $k_v$  versus  $\mathbf{X} = \{G, V, D, \Gamma, k\}$  are considered (figures 4 and 5). The effects of LPBF solidification conditions  $\mathbf{X}^s = \{G, V\}$  on  $k_v$  are shown in figure 4. The lines of best fit of the  $k_v$  data extracted from both planar and cellular interfaces are presented. On average,  $k_v$  varies between 0.5 and 0.9, which defines the microsegregation range of a Ni–Nb alloy during LPBF solidification. The  $k_v$  data are plotted

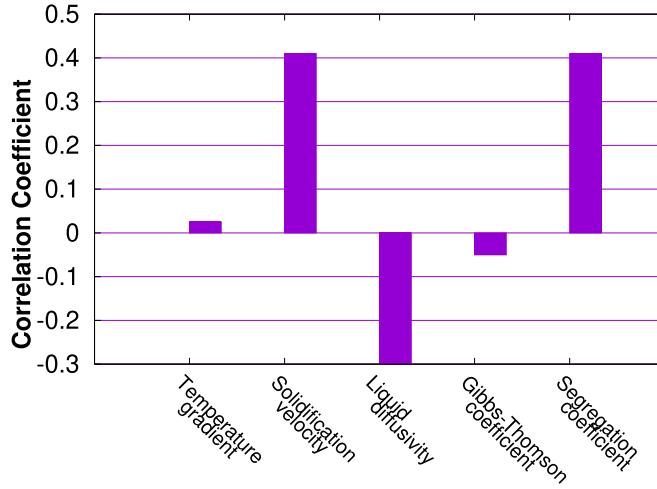


**Figure 6.** Phase-field simulations are performed for the maximum value of the velocity  $V = 2.5 \text{ m s}^{-1}$  (for which maximum deviation from the interface equilibrium is expected) to illustrate the effect of the (a) mesh size and (b) interface width on Nb microsegregation. Simulated output  $k_v$  values become grid-independent below a mesh size of 0.5 nm and interface thickness of 0.5 nm. Although not shown here, such a convergence study using the extreme values of other input parameters in table 1 resulted in a similar observation.

against the material parameters  $X^m = \{D, \Gamma, k\}$  in figure 5. Note that the line of best fit of the  $k_v$  distribution remains ‘flat’ for the changes in  $X^m = \Gamma$  (figure 5(b)) and, also for,  $X^s = G$  (figure 4(a)), meaning that there is no discernible pattern between each of these model inputs and the output. This is quantified using the correlation coefficient  $\rho_{XY}$ .

$\rho_{XY}$  defines the direction and strength in a relationship between each input and the output. In order to quantify the sources of input uncertainty on  $k_v$ , we estimate  $\rho_{XY}$  following the procedure described in section 2.5. The calculated  $\rho_{XY}$  value for each source of uncertainty is shown using a bar chart in figure 7. Note that the values of  $\rho_{XY}$  for  $G$  and  $\Gamma$  are close to 0, which signify a negligible correlation between  $k_v$  and  $X = \{G, \Gamma\}$  and explain the ‘flat’ distributions in figures 4(a) and 5(b). Thus  $k_v$  is not related in a meaningful way to the scores of  $G$  and  $\Gamma$  and it would be virtually impossible to predict  $k_v$  simply by knowing the values of  $G$  and  $\Gamma$ . On the other hand, the relationships between  $k_v$  and  $(X^s = V \text{ and } X^m = \{D, k\})$  are found to be correlated and are considered to be significant. The correlation between  $k_v$  and  $X = \{V, k\}$  is positive, meaning that  $k_v$  increases for the increase in  $X$  and vice versa. In contrast,  $D$  exhibits a negative correlation with  $k_v$ , meaning that  $k_v$  increases when  $D$  decreases and vice versa. Niobium partitioning at the solid–liquid interface increases with the increasing rate of diffusion, decreasing  $k_v$ .

A liner regression analysis is used to predict the phase-field output for any given value of  $X$ . Following the procedure described in section 2.5 and the calculation of  $\rho_{XY}$ , it is now possible to obtain the regression equations (the units of  $X$  are in SI metric) that are expressed as:



**Figure 7.** A bar chart shows the calculated values of the correlation coefficient between various sources of uncertainty and the output. The solidification velocity, liquid diffusivity, and segregation coefficient are strongly correlated with the output, while temperature gradient and Gibbs–Thomson coefficient are not.

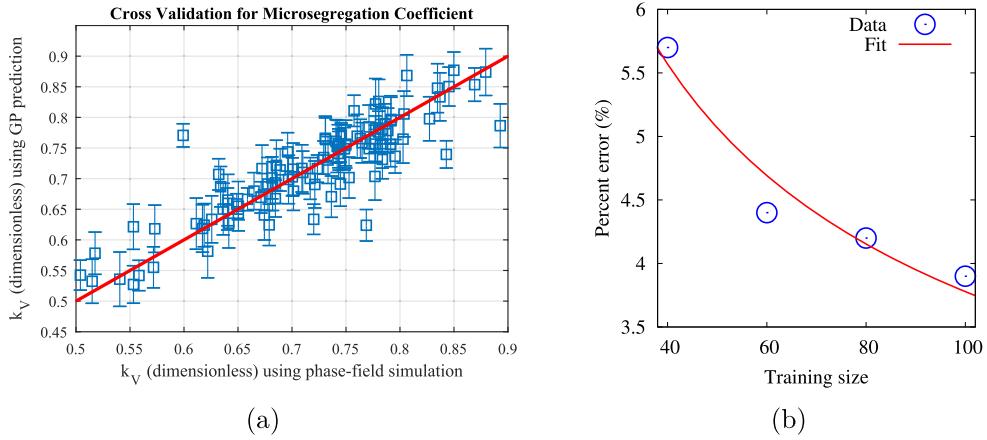
$$\begin{aligned}
 k_v &= 0.047 V + 0.6625, \\
 k_v &= -10^7 D_l + 0.77, \text{ and} \\
 k_v &= 0.78 k + 0.24.
 \end{aligned} \tag{5}$$

These equations represent the lines of least square which are essentially the ‘fitted’ lines in the scatterplots in figures 4 and 5. The slopes and intercepts of the regression lines are given in equation (5). Note the nearly ‘flat’ lines in figures 4(a) and 5(b), which signify that the  $k_v$  data points are strongly scattered and bear no correlation with the input  $X$ . Whereas in figure 5(c),  $k_v$  is aligned with a slope, meaning  $k_v$  is less scattered and a strong correlation between  $k$  and  $k_v$  exists.

### 3.4. Surrogate analysis

The above correlation and regression analyses qualitatively predict the first order relationship between  $X$  and  $Y$ . For a quantitative analysis of higher order relationships, brute force Monte Carlo simulations may be performed. However, the Monte Carlo method is computationally intensive for most computer simulation models. A surrogate model can be used instead for a quantitative assessment of the uncertainty in phase-field model inputs. Surrogate models often consist of GP or polynomial chaos emulators that provide computationally inexpensive approximations to the original computer model [13, 49]. We adopted a GP surrogate model in this work to substitute for the phase-field model and employed the framework described in [40], as implemented in [50]. A brief outline of the GP model and the procedure required to obtain the surrogate predictions are presented in [appendix](#).

The surrogate model is constructed to predict the output  $k_v$  for a given input  $X$ . Two different GP surrogate models run  $N$  computational experiments with varying values of two different input sets, one with five variable inputs  $X = \{G, V, D, \Gamma, k\}$  and the other with three inputs  $X = \{V, D, k\}$ , the parameters of which were considered significant after the correlation analysis. The corresponding surrogate model predictions and their confidence intervals are presented in figure 8(a), along with the ideal case (red line) that represents

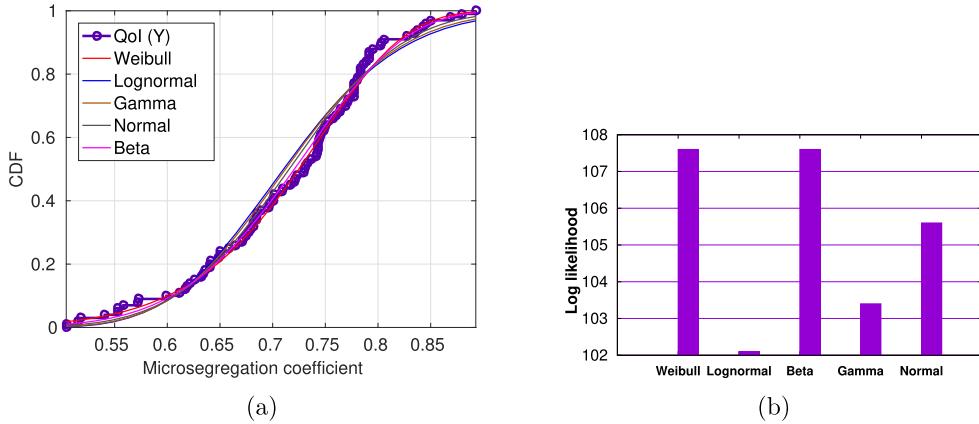


**Figure 8.** (a) A Gaussian process surrogate model predicts microsegregation using three uncertain inputs  $X = [V, D, k]$ . The horizontal axis represents the phase-field model predictions and the vertical axis represents the surrogate model predictions. The red line represents the ideal case when surrogate model predictions are in full agreement with the phase-field predictions. Each surrogate prediction is plotted with a confidence interval that represents the standard deviation around the mean. On average, the difference between the surrogate model and the phase-field predictions is satisfactory. (b) Percentage error between these predictions is shown as a function of the training data size. As expected, the difference between the training data (i.e. phase-field) and surrogate prediction decreases with increasing data size.

**Table 2.** Mean absolute predictive and percent errors of the surrogate model predictions.

Statistics	GP with three inputs	GP with five inputs
Mean absolute predictive error (MAPE)	0.0273	0.0287
Observed range in simulation	0.39	0.39
Mean absolute percent error	3.9%	4.1%

the phase-field predictions. We do not present the results of the five-parameters surrogate model, which shows a similar behavior as in figure 8(a); the statistical quantities estimated using both models are given in table 2. For a quantitative assessment, the computed mean absolute predictive (equation (4)) and percentage errors for both surrogate models are reported, the statistical measures of which indicate satisfactory performance of the surrogate models used. Note that the error metrics determined for the surrogate model with three inputs and the surrogate model with five inputs are similar, signifying that the effects of variability regarding the two additional inputs in the five-parameters model are negligible, on average. The GP surrogate model with five inputs has a slightly lower performance compared to the model with three inputs, which may be due to the over-fitting caused by more inputs compared to the model with three inputs. These results signify that the three parameters,  $V$ ,  $D$ , and  $k$ , represent the primary sources of uncertainty in the simulation of LPBF solidification process, whereas the variability in two inputs,  $G$  and  $\Gamma$ , can be neglected safely in order to reduce computational effort required for quantification.



**Figure 9.** (a) Calculated cumulative distribution of the phase-field simulated  $k_v$  population ( $N = 100$ ) is fitted against different target CDFs. (b) Maximum ‘log-likelihood’ of each fitted CDF against the empirical CDF is shown.

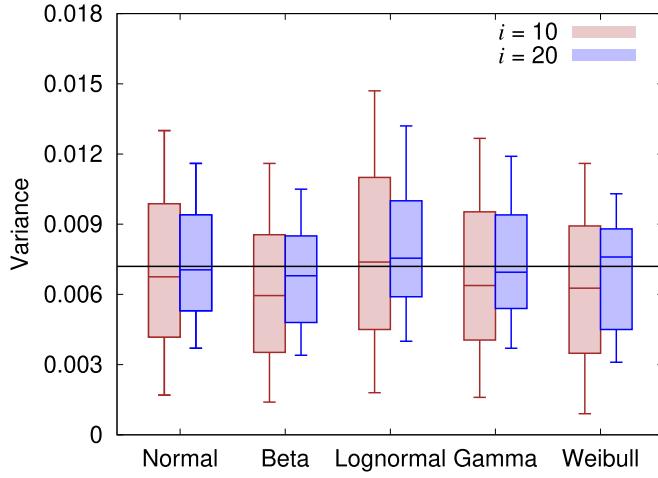
**Table 3.** Ranked ordered phase-field measurements.

Rank ( $m$ )	Output ( $k_v$ )	Relative rank ( $F$ )
1	0.504 062	0.01
2	0.514 948	0.02
3	0.517 629	0.03
4	0.540 556	0.04
5	0.553 158	0.05
..		
95	0.842 911	0.95
96	0.845 357	0.96
97	0.849 884	0.97
98	0.869 146	0.98
99	0.879 405	0.99
$N = 100$	0.892 838	1.0

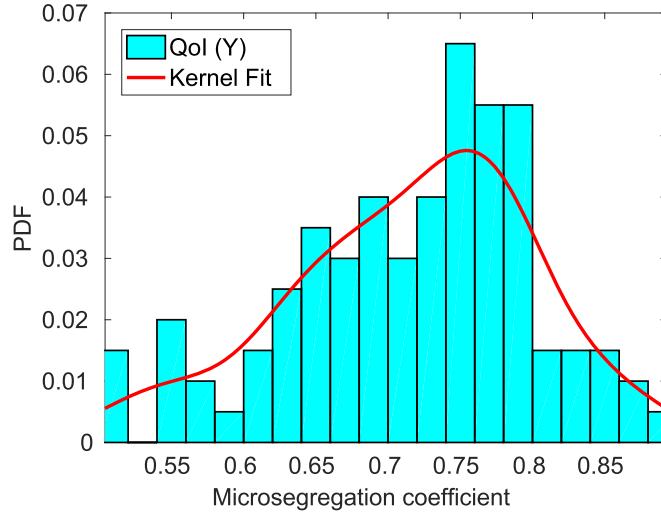
The mean absolute percent difference between the phase-field and surrogate model outputs is 3.9% (table 2). One of the primary reasons for this discrepancy can be due to a small number ( $N = 100$ ) of samples used to train the surrogate model. Figure 8(b) shows the effect of the training data size on the percent error, which can be expressed using an empirical power law of percent error =  $0.266 \times \text{size}^{-0.424}$ . In our opinion,  $N = 100$  represents the lower limit of the number of samples required to apply uncertainty analysis in materials research. More training samples are needed for a more accurate prediction of the surrogate model used in the present study.

### 3.5. Frequency analysis

We used cumulative distributions and probability densities to analyze the  $k_v$  distribution obtained from our phase-field simulations. Following the procedure described in section 2.5,  $k_v$  measurements are ranked in the ascending order, as presented in table 3, where  $F$  is the cumulative distribution. The corresponding empirical CDF is presented in figure 9(a). To gain



**Figure 10.** Phase-field simulated microsegregation distribution is fitted against different target CDFs for two samples of size  $i = 10$  (left bar) and  $i = 20$  (right bar).



**Figure 11.** Histogram of the microsegregation data and the fitted PDF using a normal kernel is illustrated.

statistical insights from the  $k_v$  data, we have fitted our empirical CDF with several theoretical CDFs that include normal, log-normal, gamma, beta, and Weibull functions. Weibull and beta functions produced maximum ‘log likelihood’ with the empirical  $k_v$  (figure 9(b)). Further, to test the effects of sampling, we have used these target CDFs to fit two different random samples of  $k_v$  of size  $i = 10$  and  $i = 20$ . We use an objective function:  $\frac{1}{N} \sum_i^N [F_i - \hat{F}_i^t]^2$ , where  $F_i$  is the empirical CDF value and  $\hat{F}_i^t$  is the corresponding fitted CDF value of a distribution  $t$ , to measure how good the average ‘fit’ is between the two. The corresponding statistical measurements are presented in figure 10. Clearly, the median of the variance for the log-normal and normal distributions is closest to the  $k_v$  distribution for  $i = 10$  and  $i = 20$

samples, respectively. However, the spread and the range of the  $k_v$  distribution are closest to a beta CDF, on average.

Figure 11 plots the histogram of the  $k_v$  data and the associated fitted PDF using a normal kernel. The peak of the histogram is close to the nominal  $k_v$  that represents the characteristics microsegregation value of an as-built Ni–Nb alloy during the LPBF process. The significance of such a PDF is twofold. The distribution can be compared with experimental measurements for a calibration of the model parameters in an inverse UQ approach (figure 1). Also, it can be used to model the subsequent heat treatment processes to aid the microstructural design in a forward UQ problem.

## 4. Discussion

### 4.1. General remarks

We present a case study illustrating how different sources of model uncertainty at the mesoscale can be identified and subsequent frequency analysis can be performed to understand the microsegregation statistics that result during the microstructure evolution in the LPBF process. We used an alloy phase-field model [26] as a reference to simulate the solidification behavior of a dilute Ni–Nb alloy (far away from Ni–Nb eutectic composition). The model explicitly tracks the solidification front, from where the Nb composition was extracted to determine the microsegregation number  $k_v$ . Note that our phase-field model is based on significant simplifications of the LPBF process and ignores the effects of several melt pool physics, such as the diffusion of heat and the convection in the liquid, on  $k_v$ . A justification for the use of these approximations to keep our phase-field simulations tractable is discussed in [23]. Ours is a baseline reference study to understand the variability of the model inputs on the microstructural features in LPBF solidification regime. Although there are process parameter ( $G, V$ ) uncertainties that originate during FEA simulations, we do not perform a multi-model-level UQ [51] in this work and focus only the influence of phase-field single-model-level UQ on microstructure statistics.

The computational requirements of the phase-field model depend on the approximations used and the domain size employed to make the simulations tractable. For example, simulating multicomponent solidification with the incorporation of additional physics such as fluid flow cost significantly more computational resources compared to simulating binary alloy solidification with reasonable approximations. Although the phase-field equations of motion (equations (1) and (2)) scale well on large parallel computer architectures, these simulations often run for days on a multi-node multi-core supercomputing framework to reach a steady state with a statistically appreciable number of microstructural features that further depend on the domain size used in the simulations.

Validation of our phase-field model against both analytical solutions and, where available, experimental measurements in the LPBF solidification limit have been reported previously [23, 52] and are therefore not repeated here. In these previous works, the concentration variation of Nb in the solidified melt pool could not be resolved during a scanning electron microscopy analysis since the beam spot size was quite large with respect to the extremely fine microsegregation features. Niobium concentration at the solid–liquid interface varies roughly between 5 wt% and 12 wt% in our simulations (figure 3). As far as we are aware, the most closely related experimental measurement of elemental distribution during LPBF is reported in figure 9(a) within [53]. These experiments however correspond to a fixed  $G$  and  $V$ , for which Nb varied between 5 wt% and 10 wt% along the traces of the solid–liquid interface in a solidified IN718 molten pool. When its binary analog is considered,

as in the present case, Nb partitioning across the interface can be more pronounced since the interactions among other alloying elements via solidification range, diffusivity, and partition coefficient are absent.

#### 4.2. Influence of parameters

The majority of our phase-field simulations resulted in planar solidification, and the remaining simulations predicted cellular solidification. In the literature [54–57], LPBF experiments were performed for a single set of laser parameters, for which the resulting solidification conditions yielded a primarily cellular microstructure. The ranges of laser parameters that provided the ranges of solidification conditions in our phase-field simulations are large when compared to an experimental study that uses a single set of laser parameter. As a result, the resultant ranges of  $G$  and  $V$  were large enough to produce combinations of cellular and planar interfaces. The simulated morphologies are consistent with the conditions of  $V_{cs}$  and  $V_{ab}$  (discussed in section 3.1); since the majority of the  $V$  samples are higher than  $V_{ab}$ , the corresponding interface morphology is planar. Such broad distributions of laser parameters and the resultant solidification conditions, on average, approximate the LPBF process and solidification maps of a generic alloy. Our results will be valuable when compared directly with uncertainties in cellular and dendritic morphologies that result for the LPBF solidification conditions. Work in these directions is currently in progress.

$G$  and  $V$  are estimated together at a particular location in the melt pool and are therefore correlated regarding the location-specific microstructural requirements during solidification. Single-track or multiple-track laser melting processes affect the solidification conditions as the neighboring melt pools interact. In reality, there is also a certain probability of finding the same value of either/both  $G$  and  $V$  at several locations in the melt pool, since these conditions primarily vary as a function of the melt pool height in a single molten pool within multiple melt pools (figure 2(c)). As a consequence, when these melt pools solidify into a microstructure, the key microstructural features may remain similar across microscopic distances. Similarly, when multiple melt pool microstructures are considered in that both of the solidification conditions may be repeated across these microstructures. Further, our analysis proves that in the high-velocity limit (or, at least where the majority of our simulation condition falls)  $G$  has only a minor contribution to the time-dependent solidification process. In this limit, on average, independent distributions of  $G$  and  $V$  can effectively assess the microstructure space that forms during LPBF. For a more accurate description of solidification, however,  $G$  and  $V$  may be shuffled together with a correlation during LHS.

Epistemic model solution uncertainties include numerical parameters such as interface width and mesh size. Convergence studies were performed in order to use appropriate values of interface width of 0.5 nm and mesh size of 0.3 nm, for which the simulated microstructures became grid independent. This way, numerical uncertainties were kept to a minimum. Model input uncertainty sources can have significant contributions to output uncertainties, and the inputs to which model and/or microstructure are sensitive are  $V$ ,  $D$ , and  $k$ . These are the aleatoric sources present in the modeling framework considered here. The temperature gradient  $G$  and the Gibbs–Thomson coefficient  $\Gamma$  do not relate to  $k_v$  in any meaningful way, and hence they can be fixed at certain specific values, e.g. mean, without considering their variability in practice (epistemic uncertainties).

The contribution of solidification conditions on the microstructure evolution in LPBF solidification regime can be ranked after the measurements of Pearson correlation. This analysis shows that  $V$  ( $\rho_{XY} = 0.42$ ) is the most important process parameter and the effects of  $G$  are negligible ( $\rho_{XY} = 0.03$ ). This explains why many researchers [52, 58] ignored the

variability in  $G$  and varied the value of  $V$  during the modeling of solidification at the high-velocity LPBF limit. Note that, in this first approach, we have only determined the contribution of an individual input on the output QoI without considering the mutual interactions among the inputs. One can use the Sobol' variance analysis [59] to estimate such mutual statistical interactions among the inputs which is given by:

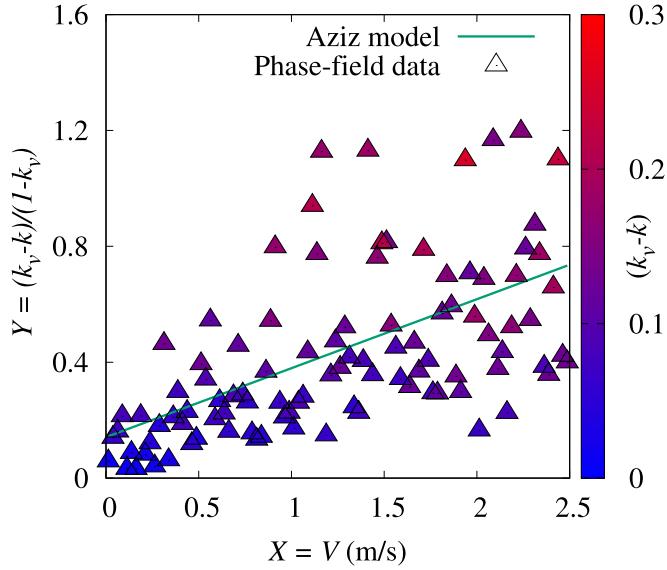
$$S_i^T = 1 - \frac{E_{X_{\sim i}}(\text{Var}_{x_i}(k_v|X_{\sim i}))}{\text{Var}(k_v)}, \quad (6)$$

where  $x_i$  is the  $i$ th input variable,  $X_{\sim i}$  is the vector of variables excluding  $x_i$ ,  $\text{Var}()$  is the variance, and  $\text{Var}_{x_i}(k_v|X_{\sim i})$  is the variance by freezing  $x_i$ . Note that the Sobol' method would require an unaffordable amount of phase-field simulations that consume significant computational time and resources. Therefore, a suitable surrogate model such as GP model [18] and polynomial chaos expansion [60] may be used to approximate the expensive phase-field calculations. Following the correlation analysis, a regression analysis was used to represent the linear relationships between each input and the output. Regression equations give an intuitive experience by providing the formula for calculating the predicted value of the output when an actual value of the input variable is known. This is our first approach, and hence we ignored statistical interactions among the input variables that can be estimated using multiple regressions [61].

#### 4.3. Surrogate modeling perspectives

A GP surrogate model was used previously to replace the FEA simulations in [17, 40]. There has been no attempt yet to use a surrogate model to quantify the process–structure–property linkages that develop during an AM process [62, 63]. Our approach to approximate the phase-field model using a GP regression model is the first step in that direction. The surrogate model predictions are satisfactory (at least for the given sample size) as they successfully replicate our phase-field predictions (figure 8). A polynomial chaos expansion [60] may be used instead of GP as a surrogate base. However, a GP surrogate model is more efficient compared to the polynomial chaos expansion in terms of capturing local structures, flexibility, and quantifying uncertainty [49].

Constructing our surrogate model was based on approximating a physics-based model using a multivariate GP model [40]. To achieve this, a number of simulations from the computationally expensive physics-based model must be generated first that uses a sampling technique such as the Latin hypercube design. One direction in the literature to improve the prediction of the surrogate model is to construct it using a combination of high-fidelity and low-fidelity simulations, rather than only using high-fidelity simulations. For example, construction of a surrogate model that uses complex high-fidelity phase-field solidification model and lower-fidelity simulations from an analytical model may train the surrogate model faster compared to generating the training points only from a high-fidelity model. Even though low-fidelity simulations on their own do not provide good predictions like the original high-fidelity physics-based model, they can be very beneficial in obtaining better surrogate model predictions than the case that only bases the surrogate model on high-fidelity simulations [64, 65]. Further, to optimally improve the predictive capability of the GP surrogate model over the domain we could develop sequential full model sampling policies based on maximizing the Kullback–Liebler divergence (equation (8)) between the current surrogate model and a surrogate model with one extra data point [66, 67]. In this sense, the process would maximize the information gained on every query to the expensive full model.



**Figure 12.** The map of Nb partitioning at the solid–liquid interface during LPBF is illustrated as a function of  $V$ . On average, the magnitude of solute trapping or the deviation from the local equilibrium (proportional to  $k_v - k$ ) increases with increasing  $V$  that is superimposed on the phase-field data using a color map. A linear regression model is fit to the data following the Aziz model [70].

#### 4.4. Microsegregation perspectives

Obtaining a quantitative basis of the microsegregation during LPBF solidification is far from straightforward. In this regime, a finite level of microsegregation is expected that is confirmed by our phase-field simulations. The estimated levels of microsegregation, however, may depend on how the anti-trapping flux term (equation (2)) is implemented in the present model. Anti-trapping flux was introduced by Karma [27] using a thin interface analysis to eliminate the artifacts due to the use of large numerical interface width during the phase-field simulation for a ‘magic’ value of  $a_t = 1/(2\sqrt{2})$ . A value of  $a_t$  larger than this may lead to the measured  $k_v$  to be lower than its equilibrium value  $k$  and therefore becomes unphysical. To a good approximation, the results presented here can be considered close to the lower limit when the phase-field model in [26, 27] was used as a reference to simulate the ranges and distribution of  $k_v$  in the LPBF regime. A more systematic work is necessary, as analyzed in [68, 69], to explore how the inclusion of  $a_t$  within our phase-field model affects the measured levels of  $k_v$  for a more accurate, physical description of solute trapping.

In experimental solidification studies, the microsegregation coefficient  $k_v$  is used to determine the characteristic interface diffusion velocity  $V_D = D/W$ , where  $D$  (equation (2)) is the diffusivity of the liquid and  $W$  (equation (1)) is the physical width of the solid–liquid interface.  $V_D$  determines the solute trapping behavior of an alloy and can be estimated by the Aziz function [70] given by:  $k_v = (k + V/V_D)/(1 + V/V_D)$ . When our phase-field  $k_v$  predictions are compared to this relationship, we obtain a solute trapping map (figure 12) where the quantity  $((k - k_v)/(1 - k_v))$  is plotted against  $V$  and a linear regression model is fit to the data to obtain the solute trapping gradient  $1/V_D = W/D$ . On average, the magnitude of solute trapping, i.e.  $(k_v - k)$ , increases with increasing  $V$  with maximum absolute uncertainty of 0.3 is predicted by our simulations (figure 12).

We numerically optimize an objective function:

$$\arg \min_{V_D} \frac{1}{N} \sum_{i=1}^N (k_v^i - k_v^i(V_D))^2 \quad \forall i \in \{1, \dots, N\}, \quad (7)$$

to find the root for  $V_D$ , which is estimated to be on the order of  $6.4 \text{ m s}^{-1}$ . When simulations were performed with the same phase-field model but with a limited input ( $i < 10$ ) of  $V$ , the  $V_D$  obtained was on the order of  $0.5 \text{ m s}^{-1}$  [23]. Clearly, the size and distribution of the model inputs affect the calculation of interfacial properties, i.e.  $V_D$ . Experiments [71] determined the  $V_D$  to be on the order of  $9.0 \text{ m s}^{-1}$  in Ni–Nb alloys. We wish to note that a generalization of the Gulliver–Scheil model [47] and Aziz model [70] could also indicate the influence of the inputs— $V$ ,  $k$ , and  $D$ —on  $k_v$  for rapid planar solidification at a much-reduced complexity compared to phase-field simulations. A comparison of the uncertainty treatments based on the analytical and phase-field methods is an interesting future research direction.

Frequency analysis of  $k_v$  distribution was performed using CDFs and PDFs. Through an analysis of CDF, it was found that the estimated CDF that fits closely to the empirical CDF was Weibull and beta, with their differences being negligible. Representation of our microsegregation dataset using such a distribution not only reduces the amount of information conveyed by the dataset, but also describes the trends in the microsegregation patterns that develop during LPBF solidification. The uncertainty of such a distribution can be determined using an entropy function  $\mathbb{H}(k_v)$ , which uses the probability  $p$  that is discretized into  $K$  states to represent the  $k_v$  distribution. The estimated  $\mathbb{H}(k_v)$  can then be used to calculate the Kullback–Leibler divergence [72] to assess the dissimilarity between two probability distributions, say  $k_{v_1}$  and  $k_{v_2}$ , using

$$\begin{aligned} \mathbb{H}(k_v) &\triangleq -\sum_{t=1}^K p(k_v = t) \log p(k_v = t) \text{ and} \\ \mathbb{KL}(k_{v_1} || k_{v_2}) &\triangleq \sum_{t=1}^K k_{v_1}^t \log \frac{k_{v_1}^t}{k_{v_2}^t}. \end{aligned} \quad (8)$$

In the literature, a Weibull distribution best described the macro-segregation behavior in the low-velocity solidification experiments and simulations [73, 74]. A lognormal distribution was shown to fit best the grain size distribution (a consequence of microsegregation) that resulted during AM experiments on Ti–6Al–4V [75]. In our calculations of a Ni–Nb alloy, estimated CDFs such as lognormal, normal and gamma are found to be either overestimating or underestimating in the tail, middle, or the top regions of the empirical CDF (figure 9(a)). When experimental measurements will be available, our analysis will be used to validate our simulations.

## 5. Summary and outlook

Our work is summarized as follows.

- Forward propagation of process and alloy parameter uncertainties through mesoscale phase-field simulations has been conducted to study the LPBF solidification process of a Ni–Nb alloy.
- Uncertainty in model parameters leads to significant variability in microstructural features. The sample size and distribution of the model parameters severely affect the QoI distribution in the LPBF solidification regime.

- Following the correlation measurements, we recommend that some less important parameters,  $G$  and  $\Gamma$ , to be fixed at certain specific values (e.g. mean), and the variability in  $V$ ,  $D$ , and  $k$  needs to be considered.
- A GP surrogate model satisfactorily approximates the phase-field predictions and hence can be used as a substitute for the phase-field method.
- A frequency analysis identifies and quantifies the microsegregation distribution in the LPBF regime that affects the built material properties at the macroscale.

AM has the potential to be the technology for the future, and the quality control in LPBF can be achieved through variation control of the QoI [6, 15]. To achieve this, a better understanding of the UQ of the LPBF process is essential. Our model and approach, in its current form, can be used as a first step towards microstructural engineering by providing users with computationally inexpensive predictions to explore the effects of model inputs on LPBF microstructures. The same approach could potentially be applied to investigate the variability on other key microstructural features, such as the dendrite arm spacing and the misorientation between dendrites. Future work will expand the proposed approach to an inverse problem on reducing the uncertainty by calibrating the model parameters using experimental measurements. The surrogate model can then be used to replace the expensive multi-scale, multi-physics phase-field model to meet the requirements of LPBF microstructures rapidly. Our ultimate goal is to build an efficient yet inexpensive framework for quantification of the datasets and distributions of the process and microstructure features in the AM solidification regime.

## Acknowledgments

The authors would like to thank the support of the National Science Foundation grant Nos. CMMI-1534534, CMMI-1663130, and DGE-1545403. RA and DA also acknowledge the support of ARL through grant No. W911NF-132-0018. Portions of this work were also supported by an Early Stage Innovations grant from NASA's Space Technology Research Grants Program, grant No. NNX15AD71G. High-throughput phase-field simulations were carried out at the Ada and Terra Texas A&M University supercomputing facilities.

## Appendix. GP surrogate model

We construct a GP surrogate model that has been detailed in [40, 76, 77]. A GP model assumes that given a finite input vector of  $n$  variables,  $\mathbf{X} = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ , the model outputs  $\mathbf{Y} = \{y_1, \dots, y_n\}$  follow a  $q$ -dimensional GP. Therefore, their joint probability distribution becomes a matrix normal distribution

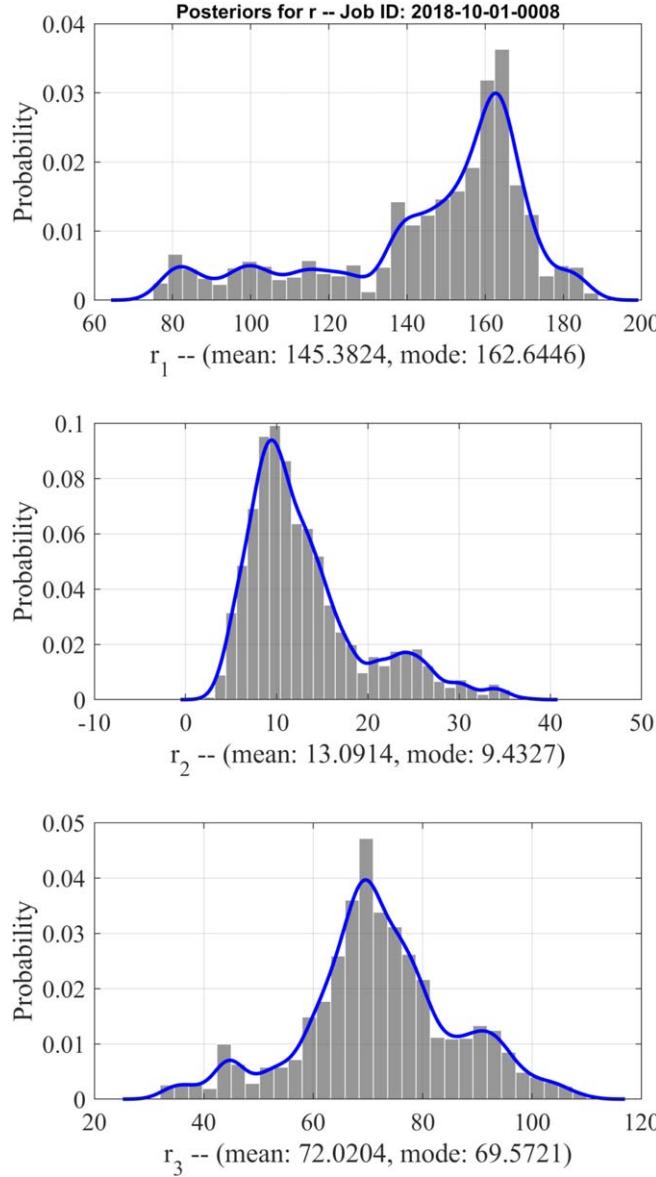
$$\mathbf{Y}|\Phi \sim \mathcal{MN}_{n,q}(\mathbf{m}, \mathbf{C}), \quad (\text{A.1})$$

where  $\mathbf{m}$  is the mean matrix and  $\mathbf{C}$  is the covariance matrix, which are fully defined by a set of hyper-parameters  $\Phi$ . The multivariate GP is expressed as

$$\mathbf{Y}|\Phi \sim \mathcal{GP}_q(m(\cdot), c(\cdot, \cdot)\Sigma), \quad (\text{A.2})$$

where  $m(\cdot)$  is the mean function,  $c(\cdot, \cdot)$  is a correlation function and  $\Sigma$  is a correlation matrix. Mean and correlation functions evaluated at  $i = j = N$  number of  $(\mathbf{X}, \mathbf{Y})$  training points are defined as:

$$m(\mathbf{x}) = \mathbf{B}^\top \mathbf{h}(\mathbf{x}) \text{ and} \quad (\text{A.3})$$



**Figure A1.** Histograms and kernel density estimates of the posterior distributions for the roughness parameters  $\mathbf{r}$  for the surrogate model with three input parameters:  $\mathbf{X} = \{D, V, k\}$ .

$$c(\mathbf{x}_i, \mathbf{x}_j) = \exp [-(\mathbf{x}_i - \mathbf{x}_j)^\top R(\mathbf{x}_i - \mathbf{x}_j)], \quad (\text{A.4})$$

where  $h$  are regression functions that maps the input space to  $m$  basis functions with regression coefficients  $\mathbf{B}$ ,  $\mathbf{x}_i = \{x_{i,1}, \dots, x_{i,n}\}_{i=1}^N$  the  $i$ th training point, and  $R = \text{diag}(\mathbf{r})$  a diagonal matrix of positive roughness parameters with  $\mathbf{r} = \{r_1, \dots, r_n\}$ .  $\mathbf{r}$  signifies the smoothness of the function. With the help of these functions, the GP model is fully defined as

$$\Phi = \{\mathbf{B}, \Sigma, \mathbf{r}\}. \quad (\text{A.5})$$

It is described in [77] that the conditional posterior distribution of  $\mathbf{Y}$  given  $\mathbf{r}$  (after integrating out  $\mathbf{B}$  and  $\Sigma$ ) is a multivariate  $q$ -dimensional  $T$  process  $f(\cdot)$  such that the resultant probability density follows a matrix-variate  $T$  distribution with  $N - m$  degrees of freedom ( $N - m$ ):

$$f(\cdot) | \mathbf{X}, \mathbf{Y}, \mathbf{r} \sim \mathcal{TP}_q(\hat{m}(\cdot), \hat{c}(\cdot, \cdot) \hat{\Sigma}, N - m). \quad (\text{A.6})$$

The  $\hat{m}$  and  $\hat{c}$  functions are defined by

$$\hat{m}(\mathbf{x}) = \hat{\mathbf{B}}^\top h(\mathbf{x}) + (\mathbf{Y} - \mathbf{H}\hat{\mathbf{B}})^\top \mathbf{A}^{-1} \mathbf{t}(\mathbf{x}) \text{ and} \quad (\text{A.7})$$

$$\begin{aligned} \hat{c}(\mathbf{x}_i, \mathbf{x}_j) &= c(\mathbf{x}_i, \mathbf{x}_j) - \mathbf{t}^\top(\mathbf{x}_i) \mathbf{A}^{-1} \mathbf{t}(\mathbf{x}_j) \\ &+ [h(\mathbf{x}_i) - \mathbf{H}^\top \mathbf{A}^{-1} \mathbf{t}(\mathbf{x}_i)^\top (\mathbf{H}^\top \mathbf{A}^{-1} \mathbf{H})^{-1} [h(\mathbf{x}_j) - \mathbf{H}^\top \mathbf{A}^{-1} \mathbf{t}(\mathbf{x}_j)]], \end{aligned} \quad (\text{A.8})$$

respectively.  $\mathbf{H}^\top = [h(\mathbf{x}_1), \dots, h(\mathbf{x}_N)]$ ,  $\mathbf{A} = c(\mathbf{x}_i, \mathbf{x}_j)$ ,  $\mathbf{t}^\top(\mathbf{x}_i) = [c(\mathbf{x}_i, \mathbf{x}_1), \dots, c(\mathbf{x}_i, \mathbf{x}_N)]$ ,  $\hat{\mathbf{B}} = (\mathbf{H}^\top \mathbf{A}^{-1} \mathbf{H})^{-1} \mathbf{H}^\top \mathbf{A}^{-1} \mathbf{Y}$ , and  $\hat{\Sigma} = (N - m)^{-1} (\mathbf{Y} - \mathbf{H}\hat{\mathbf{B}})^\top \mathbf{A}^{-1} (\mathbf{Y} - \mathbf{H}\hat{\mathbf{B}})$ . Equations (A.6)–(A.8) are used as a surrogate for the phase-field model.

To estimate the roughness parameter  $\mathbf{r}$ , the Bayesian approach is used given a positive, log-logistic prior distribution. Using the single-component Metropolis–Hastings algorithm, the posterior distributions of  $\mathbf{r}$  were generated after 20 000 iterations with 25% burn-in period and thinning every fifth sample. Two separate surrogate models were trained: one surrogate model that uses three inputs  $\mathbf{X} = \{V, D, k\}$  and another surrogate model that uses five inputs:  $\mathbf{X} = \{G, V, D, \Gamma, k\}$ . Figure A1 shows the histograms and kernel density estimates of the posterior distributions for these inputs. We observe that the posteriors are unimodal, the modes of which were used as the estimates for  $\mathbf{r}$ . At this stage, the surrogate model in equation (A.6) is fully defined, and the output of the phase-field model at any given  $\mathbf{x}_i$  can be estimated using equation (A.7). A confidence interval for this estimate can also be determined using equation (A.8).

When all the components of  $\Phi$  (equation (A.5)) are estimated, we assess the performance of the GP surrogate model through a *p-fold cross validation* (CV) to ensure that the predictions obtained using the surrogate model are close to those obtained using the phase-field model. Our CV procedure separates the training dataset  $(\mathbf{X}, \mathbf{Y})$  into  $p$  disjoint partitions, among which  $(p - 1)$  of these partitions are used to train the surrogate model, and then the predictions are made on the left-out partition using equation (A.7). These predictions are then compared with the phase-field predictions. This process is iterated  $p$  times, such that at each iteration, a different partition is left out and after  $p$  iterations all partitions have been left out only once. The comparison between surrogate and phase-field predictions is given in section 3.4.

## ORCID iDs

Supriyo Ghosh  <https://orcid.org/0000-0001-7265-5266>

## References

- [1] Frazier W E 2014 Metal additive manufacturing: a review *J. Mater. Eng. Perform.* **23** 1917–28
- [2] Herzog D, Seyda V, Wycisk E and Emmelmann C 2016 Additive manufacturing of metals *Acta Mater.* **117** 371–92

[3] Ghosh S, Ma L, Ofori-Opoku N and Guyer J E 2017 On the primary spacing and microsegregation of cellular dendrites in laser deposited Ni–Nb alloys *Modelling Simul. Mater. Sci. Eng.* **25** 065002

[4] Khairallah S A, Anderson A T, Rubenchik A and King W E 2016 Laser powder-bed fusion additive manufacturing: physics of complex melt flow and formation mechanisms of pores, spatter, and denudation zones *Acta Mater.* **108** 36–45

[5] King W E, Barth H D, Castillo V M, Gallegos G F, Gibbs J W, Hahn D E, Kamath C and Rubenchik A M 2014 Observation of keyhole-mode laser melting in laser powder-bed fusion additive manufacturing *J. Mater. Process. Technol.* **214** 2915–25

[6] King W E, Anderson A T, Ferencz R M, Hodge N E, Kamath C, Kharallah S A and Rubenchik A M 2015 Laser powder bed fusion additive manufacturing of metals: physics, computational, and materials challenges *Appl. Phys. Rev.* **2** 041304

[7] Boettinger W J, Warren J A, Beckermann C and Karma A 2002 Phase-field simulation of solidification *Annu. Rev. Mater. Res.* **32** 163–94

[8] Chen L Q 2002 Phase-field models for microstructure evolution *Annu. Rev. Mater. Res.* **32** 113–40

[9] Steinbach I 2009 Phase-field models in materials science *Modelling Simul. Mater. Sci. Eng.* **17** 073001

[10] Attari V, Ghosh S, Duong T and Arroyave R 2018 On the interfacial phase growth and vacancy evolution during accelerated electromigration in Cu/Sn/Cu microjoints *Acta Mater.* **160** 185–98

[11] Ghosh S, Mukherjee A, Abinandanan T and Bose S 2017 Particles with selective wetting affect spinodal decomposition microstructures *Phys. Chem. Chem. Phys.* **19** 15424–32

[12] Chakrabarti T and Manna S 2018 Zener pinning through coherent precipitate: a phase-field study *Comput. Mater. Sci.* **154** 84–90

[13] Allaire D and Willcox K 2010 Surrogate modeling for uncertainty assessment with application to aviation environmental system models *AIAA J.* **48** 1791–803

[14] Franco B, Ma J, Loveall B, Tapia G, Karayagiz K, Liu J, Elwany A, Arroyave R and Karaman I 2017 A sensory material approach for reducing variability in additively manufactured metal parts *Sci. Rep.* **7** 3604

[15] Ghosh S 2018 Predictive modeling of solidification during laser additive manufacturing of nickel superalloys: recent developments, future directions *Mater. Res. Express* **5** 012001

[16] Wei H, Mazumder J and DebRoy T 2015 Evolution of solidification texture during additive manufacturing *Sci. Rep.* **5** 16446

[17] Nath P, Hu Z and Mahadevan S 2018 Modeling and uncertainty quantification of material properties in additive manufacturing *2018 AIAA Non-Deterministic Approaches Conf.* p 0923

[18] Hu Z and Mahadevan S 2017 Uncertainty quantification and management in additive manufacturing: current status, needs, and opportunities *Int. J. Adv. Manuf. Technol.* **93** 2855–74

[19] Lopez F, Witherell P and Lane B 2016 Identifying uncertainty in laser powder bed fusion additive manufacturing models *J. Mech. Des.* **138** 114502

[20] Bruna-Rosso C, Demir A G, Vedani M and Previtali B 2018 Global sensitivity analyses of a selective laser melting finite element model: influential parameters identification *Int. J. Adv. Manuf. Technol.* **99** 833–43

[21] Zaeem M A, Yin H and Felicelli S D 2013 Modeling dendritic solidification of Al-3% Cu using cellular automaton and phase-field methods *Appl. Math. Modelling* **37** 3495–503

[22] Rappaz M 2016 Modeling and characterization of grain structures and defects in solidification *Curr. Opin. Solid State Mater. Sci.* **20** 37–45

[23] Ghosh S, Ofori-Opoku N and Guyer J E 2018 Simulation and analysis of  $\gamma$ -Ni cellular growth during laser powder deposition of Ni-based superalloys *Comput. Mater. Sci.* **144** 256–64

[24] Kundin J, Mushongera L and Emmerich H 2015 Phase-field modeling of microstructure formation during rapid solidification in Inconel 718 superalloy *Acta Mater.* **95** 343–56

[25] Zhang L and Steinbach I 2012 Phase-field model with finite interface dissipation: extension to multi-component multi-phase alloys *Acta Mater.* **60** 2702–10

[26] Echebarria B, Folch R, Karma A and Plapp M 2004 Quantitative phase-field model of alloy solidification *Phys. Rev. E* **70** 061604

[27] Karma A 2001 Phase-field formulation for quantitative modeling of alloy solidification *Phys. Rev. Lett.* **87** 115701

[28] Provatas N and Elder K 2011 *Phase-Field Methods in Materials Science and Engineering* (New York: Wiley)

[29] Baker H 1992 *ASM Handbook: Alloy Phase Diagrams* vol 3 (Materials Park, OH: ASM International)

[30] Knorovsky G A, Cieslak M J, Headley T J, Romig A D and Hammetter W F 1989 INCONEL 718: A solidification diagram *Metall. Trans. A* **20** 2149–58

[31] COMSOL 2011 Multiphysics Reference Guide for COMSOL 4.2

[32] Trapp J, Rubenchik A M, Guss G and Matthews M J 2017 *In situ* absorptivity measurements of metallic powders during laser powder-bed fusion additive manufacturing *Appl. Mater. Today* **9** 341–9

[33] Karayagiz K, Elwany A, Tapia G, Franco B, Johnson L, Ma J, Karaman I and Arroyave R 2018 Numerical and experimental analysis of heat distribution in the laser powder bed fusion of Ti–6Al–4V *IISE Trans.* accepted (<https://doi.org/10.1080/24725854.2018.1461964>)

[34] Block-Bolten A and Eagar T W 1984 Metal vaporization from weld pools *Metall. Trans. B* **15** 461–9

[35] Davis J and Committee A 2000 *Nickel, Cobalt, and Their Alloys: ASM Specialty Handbook* (Materials Park, OH: ASM International)

[36] Agarwal D C 2004 *Nickel and Nickel Alloys* (Hoboken, NJ: Wiley-Blackwell) ch 7, pp 217–70

[37] Ladani L, Romano J, Brindley W and Burlatsky S 2017 Effective liquid conductivity for improved simulation of thermal transport in laser beam melting powder bed technology *Additive Manuf.* **14** 13–23

[38] Roberts I A, Wang C J, Esterlein R, Stanford M and Mynors D J 2009 A three-dimensional finite element analysis of the temperature field during laser melting of metal powders in additive layer manufacturing *Int. J. Mach. Tools Manuf.* **49** 916–23

[39] Raghavan N, Dehoff R, Pannala S, Simunovic S, Kirka M, Turner J, Carlson N and Babu S S 2016 Numerical modeling of heat-transfer and the influence of process parameters on tailoring the grain morphology of IN718 in electron beam additive manufacturing *Acta Mater.* **112** 303–14

[40] Mahmoudi M, Tapia G, Karayagiz K, Franco B, Ma J, Arroyave R, Karaman I and Elwany A 2018 Multivariate calibration and experimental validation of a 3D finite element thermal model for laser powder bed fusion metal additive manufacturing *Integr. Mater. Manuf. Innov.* **7** 116–35

[41] Yuan L and Lee P D 2010 Dendritic solidification under natural and forced convection in binary alloys: 2D versus 3D simulation *Modelling Simul. Mater. Sci. Eng.* **18** 055008

[42] Olsson A, Sandberg G and Dahlblom O 2003 On Latin hypercube sampling for structural reliability analysis *Struct. Saf.* **25** 47–68

[43] Montgomery D and Runger G 2014 *Applied Statistics and Probability for Engineers* (New York: Wiley)

[44] Devore J 1987 *Probability and Statistics for Engineering and the Sciences* (Boston, MA: Brooks/Cole)

[45] Kennedy M C and O'Hagan A 2001 Bayesian calibration of computer models *J. R. Stat. Soc. B* **63** 425–64

[46] MATLAB 2018 Version R2018a (Natick, MA: The MathWorks Inc.)

[47] Rappaz M and Dantzig J A 2009 *Solidification, Engineering Sciences* (Lausanne: EPFL Press)

[48] Mullins W W and Sekerka R F 1964 Stability of a planar interface during solidification of a dilute binary alloy *J. Appl. Phys.* **35** 444–51

[49] O'Hagan A 2013 Polynomial chaos: a tutorial and critique from a statistician's perspective *SIAM/ASA J. Uncertain. Quantif.* **20** 1–20

[50] Mahmoudi M and Tapia G 2018 Multivariate Statistical Calibration of Computer Models <https://github.com/mahmoudi-tapia/MVcalibration>

[51] Sankararaman S and Mahadevan S 2012 Likelihood-based approach to multidisciplinary analysis under uncertainty *J. Mech. Des.* **134** 031008

[52] Keller T et al 2017 Application of finite element, phase-field, and CALPHAD-based methods to additive manufacturing of Ni-based superalloys *Acta Mater.* **139** 244–53

[53] Acharya R, Sharon J A and Staroselsky A 2017 Prediction of microstructure in laser powder bed fusion process *Acta Mater.* **124** 360–71

[54] Ghosh S, Ma L, Levine L E, Ricker R E, Stoudt M R, Heigel J C and Guyer J E 2018 Single-track melt-pool measurements and microstructures in Inconel 625 *JOM* **70** 1011–16

[55] Ghosh S, Stoudt M R, Levine L E and Guyer J E 2018 Formation of Nb-rich droplets in laser deposited Ni-matrix microstructures *Scr. Mater.* **146** 36–40

[56] Vrancken B, Thijs L, Kruth J-P and Humbeeck J Van 2014 Microstructure and mechanical properties of a novel  $\beta$  titanium metallic composite by selective laser melting *Acta Mater.* **68** 150–8

[57] Tao P, Li H, Huang B, Hu Q, Gong S and Xu Q 2019 The crystal growth, intercellular spacing and microsegregation of selective laser melted Inconel 718 superalloy *Vacuum* **159** 382–90

[58] Boettinger W J and Warren J A 1999 Simulation of the cell to plane front transition during directional solidification at high velocity *J. Cryst. Growth* **200** 583–91

[59] Li C and Mahadevan S 2016 An efficient modularized sample-based method to estimate the first-order Sobol' index *Reliab. Eng. Syst. Saf.* **153** 110–21

[60] Tapia G, King W E, Arroyave R, Johnson L, Karaman I and Elwany A Validation of a laser-based powder bed fusion thermal model via uncertainty propagation and generalized polynomial chaos expansions *J. Manuf. Sci. Eng.* (<https://doi.org/10.1115/1.4041179>)

[61] Aiken L, West S and Reno R 1991 *Multiple Regression: Testing and Interpreting Interactions* (London: Sage)

[62] Popova E, Rodgers T M, Gong X, Cecen A, Madison J D and Kalidindi S R 2017 Process-structure linkages using a data science approach: application to simulated additive manufacturing data *Integr. Mater. Manuf. Innov.* **6** 54–68

[63] Jung J, Yoon J I, Park H K, Kim J Y and Kim H S 2019 An efficient machine learning approach to establish structure-property linkages *Comput. Mater. Sci.* **156** 17–25

[64] Kennedy M C and O'Hagan A 2000 Predicting the output from a complex computer code when fast approximations are available *Biometrika* **87** 1–13

[65] Tuo R, Wu C J and Yu D 2014 Surrogate modeling of computer experiments with different mesh densities *Technometrics* **56** 372–80

[66] Huan X and Marzouk Y 2014 Gradient-based stochastic optimization methods in Bayesian experimental design *Int. J. Uncertain. Quantif.* **4** 479–510

[67] Ghoreishi S F, Molkeri A, Srivastava A, Arroyave R and Allaire D 2018 Multi-information source fusion and optimization to realize ICME: application to dual-phase materials *J. Mech. Des.* **140** 111409

[68] Mullins A M, Rosam J and Jimack P K 2010 Solute trapping and the effects of anti-trapping currents on phase-field models of coupled thermo-solutal solidification *J. Cryst. Growth* **312** 1891–7

[69] Ohno M and Matsuura K 2009 Quantitative phase-field modeling for dilute alloy solidification involving diffusion in the solid *Phys. Rev. E* **79** 031603

[70] Aziz M J 1982 Model for solute redistribution during rapid solidification *J. Appl. Phys.* **53** 1158–68

[71] Tian Y, Muñiz-Lerma J and Brochu M 2017 Nickel-based superalloy microstructure obtained by pulsed laser powder bed fusion *Mater. Charact.* **131** 306–15

[72] Murphy K and Bach F 2012 *Machine Learning: A Probabilistic Perspective, Adaptive Computation and Machi* (Cambridge, MA: MIT Press)

[73] Voller V and Vušanović I 2014 Frequency analysis of macrosegregation measurements and simulations *Int. J. Heat Mass Transfer* **79** 468–71

[74] Fezi K, Plotkowski A and Krane M J 2016 A metric for the quantification of macrosegregation during alloy solidification *Metall. Mater. Trans. A* **47** 2940–51

[75] Loughnane G 2015 A framework for uncertainty quantification in microstructural characterization with application to additive manufacturing of Ti–6Al–4V *PhD Thesis* Wright State University

[76] Rasmussen C and Williams C 2006 *Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning Series)* (Cambridge, MA: MIT Press)

[77] Conti S and O'Hagan A 2010 Bayesian emulation of complex multi-output and dynamic computer models *J. Stat. Plan. Inference* **140** 640–51