Application of a Chained-ANN for Learning the Process-Structure Mapping in $Mg_2Si_xSn_{1-x}$ Spinodal Decomposition

Grayson H. Harrington[1], Conlain Kelly[1], Vahid Attari[2], Raymundo Arroyave[2], Surya R. Kalidindi[1,3*]

[1]Georgia Institute of Technology, School of Computational Science Engineering, Atlanta, GA 30332
[2]Texas A&M University, Materials Science and Engineering Department, College Station, TX 77843
[3]Georgia Institute of Technology, George W. Woodruff School of Mechanical Engineering, Atlanta, GA 30332

*Correspondence: surya.kalidindi@me.gatech.edu

Abstract

This work establishes a reliable and accurate materials process-structure (PS) surrogate model that maps an 18-dimensional process parameter input domain to a high dimensional space of single and dual-phase microstructures. This was accomplished by employing the Materials Knowledge Systems (MKS) framework (includes microstructure quantification via two-point statistics and dimensionality reduction using principal components analysis) for the feature engineering of the microstructures, and subsequently constructing a chained-artificial neural network (ANN) to learn the complex nonlinear mappings between the high-dimensional input domain and the MKS-derived low-dimensional representation of the corresponding microstructure space (includes both homogeneous and heterogeneous microstructures). The benefits of this workflow are demonstrated on a collection of ~10,000 final microstructures obtained from chemo-mechanical spinodal decomposition phase-field simulations in the $Mg_2Si_xSn_{1-x}$ material system. Specifically, it is shown that the complex phase-field process-structure relationships for the selected case study can be captured in a robust model with only 742 fittable parameters.

Keywords: Machine Learning**,** Chained-ANN**,** Materials Informatics**,** Process-Structure Surrogate Model**,** Phase-Field Surrogate Model

1. Introduction

Physics-based multiscale materials simulation tools [1–5] offer tremendous promise for the optimal design of novel materials with enhanced properties. However, their high computational cost hinders their broader adoption and usage by industry in practical materials design efforts [2,6]. Materials informatics tools [6–8] can potentially bridge this gap by training low-computational cost process-structure-property (PSP) surrogate models of the expensive physics-based multiscale materials simulations. Once trained[1], the surrogate model can be used in place of the expensive physics-based model for efficiently exploring material process design solutions. Among the surrogate PSP models, the construction of process-structure (PS) models is often more difficult than that of structure-property (SP) models due to the high-dimensionality of both the material manufacturing processing paths and the resultant microstructures (note that SP linkages often map microstructures to a single property value).

The PS model building efforts encounter two main challenges: (1) adequate microstructure representation (i.e., the quantification of the salient features of the resultant microstructure) and (2) ensuring sufficient expressivity of the model building approach (i.e., ability to capture the highly nonlinear mappings involved) while avoiding overfits. Prior PS surrogate modeling techniques have largely circumvented these challenges by limiting attention to purposefully simplified mappings between process and structure [9–15]. For example, Sarkar et al. [9] recently established a surrogate model for the process-structure mapping of ZrO_2-toughened Al_2O_3 ceramics subjected to sintering heat treatments. In this work, only the sinter time, sinter temperature, ceramic sinter-aid content, and reinforcement addition content were used to predict final sinter density and grain size. Here the process space is reduced to only four distinct variable parameters and the microstructure representation is drastically reduced to two features. It is noteworthy that this mapping function could be established successfully using a simple polynomial regression. In another study, Tapia et al. [10] demonstrated the feasibility of using Gaussian Process Regression (GPR) to model secondary particle nucleation and growth during heat treatment of NiTi shape-memory alloys. The inputs to their GPR model included the temperature and duration of heat treatment as well as the initial nickel composition, while the output was the final nickel content in the metal matrix. Note that the use of a limited number of inputs and outputs in both studies described above leads to their limited utility. This is not only because of the

[1] This usually incurs a one-time high computational cost.

relatively small input domains over which the surrogate models have been trained, but also because they have filtered out much of the important morphological information in the microstructures. Specifically, the rich details of the microstructure morphology cannot be accurately reconstructed from the simple outputs predicted by these models.

Comprehensive but compact feature engineering that captures the many rich details of the microstructure needs to be combined with expressive modeling strategies to address the PS model building challenge described above. In this paper, we address this need by leveraging the feature engineering of the Materials Knowledge Systems (MKS) [6,8,16] framework with a chained-artificial neural network (ANN). The benefits of the MKS feature engineering have been demonstrated in numerous prior PSP surrogate modeling efforts [8,11,12,17–19]. This approach involves the computation of 2-point spatial correlations for the statistical quantification of the microstructure, followed by principal component analysis (PCA) to obtain a low-dimensional representation of the spatial correlations. In this work, we will employ chained ANNs to capture the highly complex and nonlinear mappings between the high-dimensional process space and the MKS-feature engineered microstructure space. Specifically, our goal is to train an ANN to first identify the class of the output microstructure (i.e., as a single-phase or a two-phase microstructure), and subsequently chain an additional ANN for the prediction of the two-phase microstructures. The combination of the MKS framework with the high expressivity of the chained-ANN is expected to result in efficient learning of the complex PS mappings. It will be demonstrated that the proposed strategy produces robust PS models with a relatively low number of fitting (trained) parameters.

We will demonstrate the benefits of this PS modeling strategy described above on a large dataset aggregated from phase-field simulations. Our interest will be restricted to the prediction of the quasi-equilibrium microstructures close to completion of the spinodal decomposition phase-transformation (i.e., we will not consider the subsequent coarsening phenomena that typically occurs at a drastically slower rate)[2]. We aim to consider a large number of user-specified inputs to the commonly used phase-field models (these typically specify the initial conditions and the physical parameters controlling microstructure evolution) as inputs to our surrogate models, while the range of the output would cover all possible multi-phase microstructures in the selected

[2] Surrogates can also be formulated to predict microstructure evolution (i.e., predict the microstructure at the next time step based on known microstructures at previous time steps and known processing conditions) [8,11,17,30,31].

material system. Phase-field surrogate models covering such large input domains and output ranges have not yet been demonstrated in prior literature and are highly desirable to support future materials and process design efforts.

The application of machine learning approaches on phase-field datasets is indeed an active area of research (e.g., [20–27]). A number of these studies have focused on microstructure classification and optimization [20,21,26,27] as well as SP surrogates [22–25]. Here, we focus specifically on the extraction of PS surrogates from phase-field simulation data. Much like the earlier discussion, prior efforts on building phase-field PS surrogates has also been limited to process and microstructure spaces of low complexity. For example, Farizhandi et al. [28] developed an inverse process-structure surrogate via a deep neural network (DNN) to predict phase-field simulation processing conditions from microstructure input. The simulation processing conditions included only the chemical composition, temperature, and simulation duration. Notably, due to the lack of microstructure feature engineering and the choice of model building strategy, ~66,000,000 parameters were fit while training on 11,500 phase-field simulated microstructures. As another example, Herman et al. [29] created a phase-field-informed surrogate model to predict microstructure morphologies of physical vapor deposition (PVD) simulations. The dataset for building this surrogate model was populated by varying the phase fraction and deposition rate of the deposited material. Importantly, the MKS framework was utilized for microstructure feature engineering and a polynomial chaos expansion was used to formulate the PS surrogate. Herman et al. note that the PVD phase-field model was simplified by considering only phase fraction and deposition rate. Similarly, Yabansu et al. [11] utilized the MKS framework to predict the phase-field simulated microstructures during three-phase spinodal decomposition of the Al-Ag-Cu system as a function of two process parameters (temperature and time) using a simple polynomial regression. Alternatively, still using the MKS framework for feature engineering, Hu et al. [30] and Montes et al. [31] developed PS surrogates using recurrent neural networks (RNN) to learn the spinodal decomposition microstructural evolution of a binary alloy. The RNNs built in these studies utilized ~150,000 model-fit parameters trained on 5,000 microstructure evolution sequences. For the previously reported phase-field PS surrogates built using ANNs [28,30,31], the number of model-fit parameters utilized is significantly larger than the number of training datapoints; this can potentially result in overfitting of the model.

Our goal in this paper is to dramatically expand the dimensionality and sizes of the input domains and output ranges of the surrogate models trained on phase-field simulations. This will be accomplished by leveraging the inherent strengths of both the MKS framework (for feature engineering) and ANNs (high model expressivity), while utilizing a relatively small number of fittable model parameters. Specifically, we will establish PS surrogates trained on phase-field simulations in the $Mg_2Si_xSn_{1-x}$ material system for which a suitable dataset has been aggregated and shared in a public repository known as the Open Phase-field Microstructure Database (OPMD) [32,33]. The dataset covers a total of eighteen simulation inputs (i.e., thermodynamic, kinetic, elastic, and physical parameters) and the corresponding output microstructure space. Notably, the phase-field simulated microstructures contained in this database exhibit extremely large diversity in phase volume fractions as well as size and shape distributions of the constituent phase regions. In prior work using this dataset, Kunselman et al. [26,27] demonstrated the utility of semi-supervised and unsupervised classification techniques to automatically detect different classes of two-phase microstructures with high confidence. In another study [34], a multimodal machine learning approach was applied to predict selected characteristics of the phase-field simulated final microstructures (such as min/max composition and phase area fractions) while using as inputs both the phase-field simulation inputs and output microstructures. While these studies on the OPMD demonstrate the potential of modern machine learning strategies applied to understanding complex microstructure spaces, they do not establish low-computational cost surrogates needed to support materials innovation. The proposed combination of the MKS framework with a chained-ANN allows the formulation of robust PS surrogates, with relatively few model fit parameters, which successfully capture the complex mappings between the 18-dimensional process space and its corresponding, much higher dimensional, microstructure space.

2. Background

2.1 Microstructure Quantification

As already stated, the MKS framework will be utilized in this work for the low-dimensional representation of the microstructure. This approach utilizes (1) 2-point spatial correlations for rigorous statistical quantification of the microstructure, and (2) PCA for their reduced-order representations. This framework starts with a digital representation of the microstructure as a

multi-dimensional array m_s^h where h indexes the material local state, s indexes spatial bins (i.e., pixels or voxels) [7,35,36], and the value of m_s^h reflects the volume fraction of material local state h present in the spatial bin s. The material local state can be defined to include any attributes needed to describe the physical properties of the material at the scale of the spatial bins. Often, the material local state is used to denote attributes such as phase identifiers, chemical compositions, and lattice orientations. Microstructure statistics capturing important morphological information can be computed as 2-point spatial correlations, $f_{\boldsymbol{r}}^{hh'}$, reflecting the joint probability of finding material local states h and h' separated by a vector indexed by $\boldsymbol{r}$. Mathematically, these are expressed as (assuming the microstructure exhibits periodic boundary conditions) [37]

$$f_{\boldsymbol{r}}^{hh'} = \frac{\sum_{\boldsymbol{s} \in S} m_s^h \, m_{s+\boldsymbol{r}}^{h'}}{c(S)} \tag{1}$$

where S and $c(S)$ denote the set of all spatial bins in the microstructure and the cardinality of S, respectively. The sets of spatial correlations where $h = h'$ and $h \neq h'$ are known as autocorrelations and cross correlations, respectively. A single set of spatial correlations for fixed *h* and *h'* will have the same dimensionality (i.e., same number of spatial statistics) as the original microstructure m_s^h. One can compute n^2 sets of spatial correlations, where *n* is the number of local states considered. Niezgoda et al. [37] have shown that only *(n-1)* sets of these are independent. For a two-phase microstructure, a single auto correlation is adequate to capture all the independent 2-point spatial correlations.

The number of microstructure statistics tend to be far too many for practical surrogate model building and PCA has been shown to be quite effective in producing data-driven low-dimensional features from collections of computed spatial correlations [11,17,38]. PCA performs a distance-preserving rotational transformation that maximizes the capture of variance in the minimum number of terms for a given dataset (in the present context, this refers to a collection of spatial correlations computed on an ensemble of microstructures). The transformed axes are known as PC basis and the transformed data coordinates are referred to as PC scores. PCA has been found to offer many advantages in building PSP surrogates for a broad range of materials design applications. First, it captures the salient microstructure information in a small number of PC scores, ordered by relative importance (i.e., how much variation they capture). Second, it allows for an easy reconstruction of the 2-point spatial correlations up to a truncation error; this error can

be reduced by including more PC scores. Additionally, methods to reconstruct microstructures from the 2-point spatial correlations have already been discussed [39–42]. Third, these protocols have been found to be applicable across a broad range of material classes and length scales [6,11,13,18,19,38], supporting systematic integration of materials knowledge extracted from diverse applications in multiscale materials design efforts.

2.2 Surrogate PSP Model

Reduced-order PSP models can be built using a variety of surrogate model building strategies (e.g., [9–11,13,17,19,24,28–31]). ANNs are particularly attractive for capturing the highly complex nonlinear mappings present in materials phenomena. The most basic ANN architectures connect the input to the output using multilayer perceptrons (MLPs) in a fully connected feedforward network. A single perceptron (see the callout in Fig. 1) applies a linear transformation on the inputs, followed by a nonlinear activation function (e.g., ReLU, sigmoid, tanh) [43]. The fittable model parameters (principally, the weights and biases) are tuned through backpropagation in which prediction error, defined by a loss function, is minimized via stochastic gradient descent [44]. The total number of fittable parameters, P, in an MLP can be computed as

$$P = \sum_{l=1}^{L} n_l \, (n_{l-1} + 1) \tag{2}$$

where L is the total number of layers (excluding the input layer) and n_l is the number of nodes in layer l.

One drawback of ANNs is that they are prone to overfitting, especially in problems with a small training dataset (e.g., more fittable parameters compared to the number of training data points). A common practice is to partition the full dataset randomly into a training set and a test set such that they both exhibit similar output value distributions. Once the test set is identified, it is not exposed to the training process in any manner. Additionally, a small portion of the training set is often set aside for validation, which is aimed at mitigating model overfit. Model parameter fitting is performed using the training set and the validation set is used to determine hyperparameters such as model architecture or to implement early stopping strategies during the minimization of the loss function. Note that the test set is used exclusively to measure model

accuracy only after all model training/tuning is complete, ensuring an unbiased estimate of the model accuracy on previously unseen data.

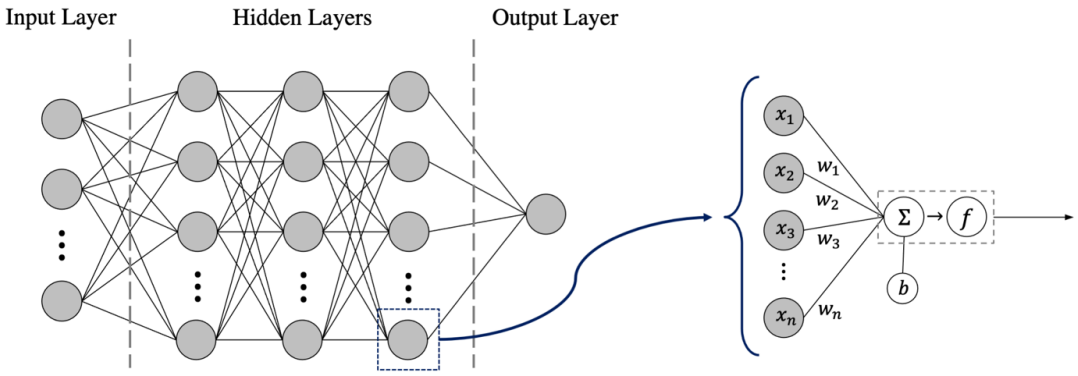

Figure 1. Schematic of a fully connected MLP for a single output mapping. The callout shows a single perceptron as a weighted linear combination of all inputs to the node with a bias, followed by the application of a nonlinear activation function.

An ANN can be designed to serve either as a classifier (predicts discrete classes) or a regressor (predicts continuous real values). Different error metrics are needed for these different applications. Classification model accuracy can be assessed using precision (P), recall (R), and $f1$-scores, defined as

$$P = \frac{TP}{TP + FP} \tag{3}$$

$$R = \frac{TP}{TP + FN} \tag{4}$$

$$f1 = 2\frac{P\,R}{P + R} \tag{5}$$

where TP is the count of true-positive predictions, FP is the count of false-positive predictions, and FN is the count of false-negative predictions. As model accuracy increases, all three metrics (P, R and $f1$) approach unity.

Regression model accuracy can be quantified using normalized mean absolute error $nMAE$) defined as

$$nMAE = \frac{\sum_{i=1}^{N} |y^i - \tilde{y}^i|}{\sum_{i=1}^{N} |y^i|} \quad (6)$$

where y^i and $\tilde{y}^i$ denote, respectively, the true output (i.e., ground-truth) and the predicted model output for the i^{th} datapoint, and N is the total number of test samples.

3. Spinodal Decomposition Process-Structure Surrogate

3.1 The OPMD Dataset

The OPMD [32] is a collection of consistently computed phase-field microstructures of the $Mg_2Si_xSn_{1-x}$ thermoelectric alloy during isothermal annealing. The $Mg_2Si_xSn_{1-x}$ pseudo-binary material system is an important thermoelectric alternative to the common, environmentally harmful, thermoelectric materials in current use (e.g., PbTe, Bi_2Te_3, and skutterudites) [45–47]. This material system is well known for its miscibility gap in which spinodal decomposition (spontaneous phase separation) is chemically favorable. Such solid-state reactions are also influenced by the interfacial and elastic energy contributions to the bulk free energy of the system. While phase separation is promoted by chemical driving forces, the elastic driving forces often suppress the miscibility gap and promote phase dissolution or homogenization [45]. These competing driving forces (chemical and elastic) can be modelled in a fully parameterized elasto-chemical phase-field simulation in which the effect of the various physical parameters on $Mg_2Si_xSn_{1-x}$ microstructure can be studied. The derivation and use of the phase-field modeling for the spinodal decomposition in $Mg_2Si_xSn_{1-x}$ have been discussed extensively by Attari et al. [33].

Currently, the OPMD contains 10,000 phase-field produced microstructure evolution (presented as time-series) simulations along with all of the relevant values of the input parameters involved. More specifically, these simulations utilize a total of eighteen input parameters describing the thermodynamic, kinetic, elastic, and physical properties of the $Mg_2Si_xSn_{1-x}$ material system. The ranges for these input parameters are summarized in Table 1. The CALPHAD parameters for $Mg_2Si_xSn_{1-x}$ solid solution (${}^0a^{ss}, {}^0b^{ss}, {}^1a^{ss}$) and for the liquid phase (${}^0a^{liq}$, ${}^0b^{liq}, {}^1a^{liq}$) describe the interactions between the constituents beyond the ideal mixtures. These parameters were used to estimate the molar Gibbs free energy of mixing for the pseudo-binary system, which was used as an input to the phase-field simulations. The kinetic parameters shown

in Table 1 control the mass diffusion rates, while the elastic constants ($C_{11}^{Mg_2Sn}$, $C_{12}^{Mg_2Sn}$, $C_{44}^{Mg_2Sn}$, $C_{11}^{Mg_2Si}$, $C_{12}^{Mg_2Si}$, $C_{44}^{Mg_2Si}$) control the elastic driving forces. The range for the Stress-Free Transformation Strain ($SFTS$) was estimated using both calculated and experimental lattice parameters in the literature for Mg_2Sn and Mg_2Si. The ranges for the molar volumes of Mg_2Sn and Mg_2Si were based on cell volume calculations and experimental measurements, and the alloy composition was sampled in the range $0.3 \leq x \leq 0.5$ to ensure that the phase-field simulations were executed within the chemically unstable miscibility gap for the $Mg_2Si_xSn_{1-x}$ system at 720°C [33,45].

The simulation cell (384x384 voxels and 350x350 nm in dimension) was initialized with random noise surrounding the initial composition (defined by the Si stoichiometry, x) and each simulation was executed with an isothermal temperature of 720 °C. At each time step of the simulation, the microstructure is defined as a two-dimensional spatial field of size 384×384 describing the Si stoichiometry (expressed in the range 0 to 1) at each pixel. We extracted only the microstructure images corresponding to the 15,000th time step[3] of the phase-field simulations. The ensemble of the extracted microstructures exhibited a large amount of morphological variation (see Fig. 2), indicating a highly nonlinear mapping with the 18-dimensional input domain. It is seen that the simulation output exhibited single-phase, two-phase, and three-phase microstructures. The single-phase microstructures are those that did not phase separate by the 15,000th timestep in the phase-field simulations. The two-phase microstructures represent the cases where the parent phase ($Mg_2Si_xSn_{1-x}$ solid solution) fully decomposed into two distinct phases (Mg_2Si and Mg_2Sn) and the three-phase microstructures represent the cases where the parent phase only partially separated by the 15,000th timestep.

[3] The 15,000th time step corresponded to fully decomposed microstructures for most of the simulations. However, the phase separation was sluggish in a few cases and produced three-phase microstructures. As noted later, the small number of three-phase microstructures found in the extracted dataset were excluded from the study. Therefore, the microstructures obtained at the 15,000th time step served as the final microstructures for our study.

Table 1: Summary of the ranges of the phase-field simulation parameters for the dataset used in this work. "ss" and "liq" denote $Mg_2Si_xSn_{1-x}$ solid solution and liquid phases, respectively. SFTS denotes Stress-Free Transformation Strain [33].

	Parameter	Unit	Lower Bound	Upper Bound
CALPHAD Parameters	${}^0a^{ss}$	J mol^{-1}	6824.89	20474.69
	${}^0b^{ss}$	J mol^{-1} K^{-1}	3.67	11.02
	${}^1a^{ss}$	J mol^{-1}	-5208.34	-1736.11
	${}^0a^{liq}$	J mol^{-1}	43550.19	130650.57
	${}^0b^{liq}$	J mol^{-1} K^{-1}	-86.03	-28.68
	${}^1a^{liq}$	J mol^{-1}	3314.80	9944.41
Kinetic Parameters	Interface mobility (M)	m^2 s^{-1} J^{-1}	$10^{-20}/(RT)$	$10^{-18}/(RT)$
	Gradient energy coefficient ($\varkappa$)	J m^{-2}	2.0×10^{-26}	2.0×10^{-24}
Elasticity Parameters	SFTS (ϵ^T)	-	-0.02	0.02
	$C_{11}^{Mg_2Sn}$	GPa	68.30	83.71
	$C_{12}^{Mg_2Sn}$	Gpa	17.68	39.79
	$C_{44}^{Mg_2Sn}$	Gpa	16.03	41.94
	$C_{11}^{Mg_2Si}$	Gpa	114.07	126.00
	$C_{12}^{Mg_2Si}$	Gpa	19.56	26.00
	$C_{44}^{Mg_2Si}$	Gpa	33.32	58.20
Physical Parameters	Molar Volume $\left(V_m^{Mg_2Sn}\right)$	m^3 mol^{-1}	4.73×10^{-5}	6.38×10^{-5}
	Molar Volume $\left(V_m^{Mg_2Si}\right)$	m^3 mol^{-1}	3.95×10^{-5}	5.33×10^{-5}
	Alloy composition (x)	mol	0.3	0.5

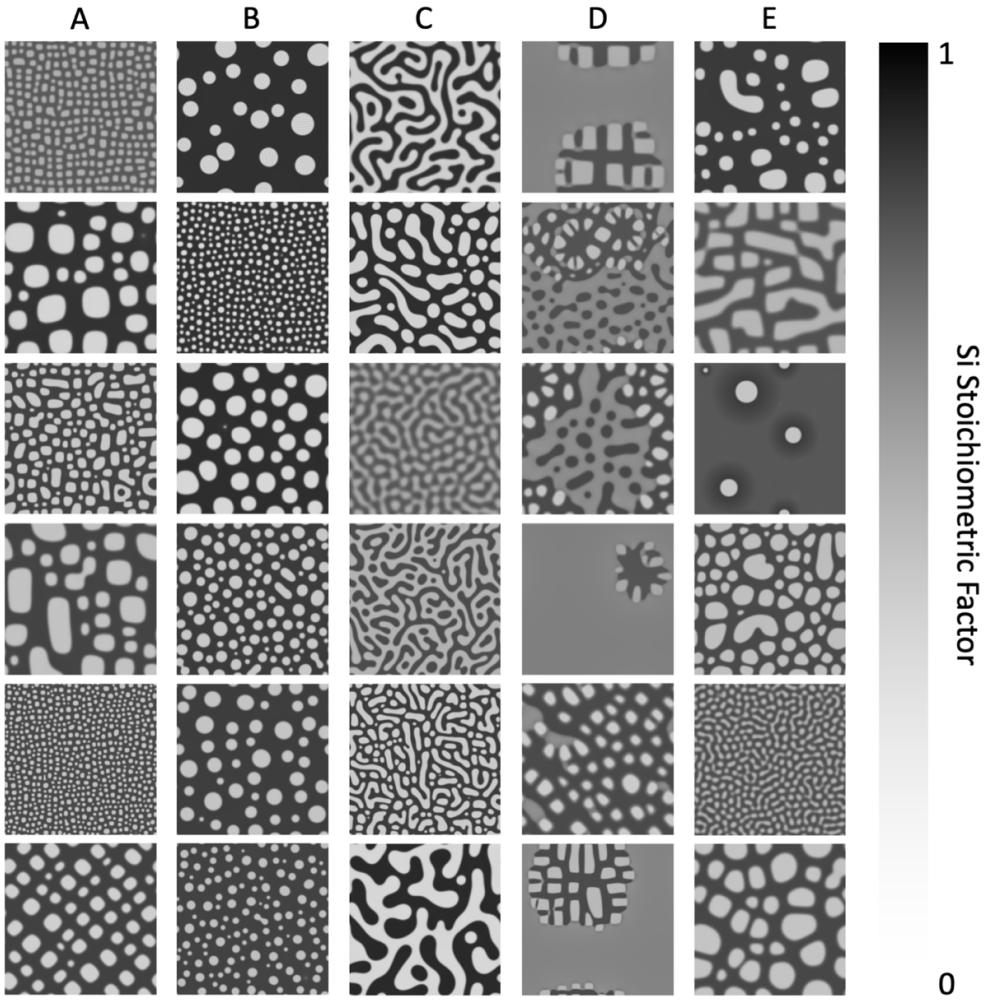

Figure 2. Example microstructures from the OPMD dataset where each column corresponds to A) cuboidal Mg_2Si nanoprecipitates, B) spherical Mg_2Si nanoprecipitates, C) intertwined Mg_2Si and Mg_2Sn grains, D) partially decomposed $Mg_2Si_xSn_{1-x}$, and E) miscellaneous microstructures

3.2 Segmentation of Simulated Microstructures

The phase-field model outputs spatial field (384×384) predictions for the Si stoichiometric factor over the entire simulated region. Although the phase-field predicted Si stoichiometry exhibits values in the continuous range of (0,1), the histograms showed only one, two or three strong peaks suggesting that the predicted microstructures indeed belong to the three classes of microstructures (i.e., single-, dual-, or triple-phase) mentioned earlier. In order to apply the microstructure quantification framework presented in Section 2.1, each pixel in each phase-field predicted spatial field needs to be labelled with one of the three possible phase labels (i.e., Mg_2Sn,

$Mg_2Si_xSn_{1-x}$, and Mg_2Si). This process (illustrated in Fig. 3 using examples from all three classes of predicted microstructures seen in the OPMD), where a continuous range is quantized, is generally referred to as segmentation and was performed via histogram analysis and Otsu thresholding [48]. When only one peak is observed, all pixels in the microstructure were classified as $Mg_2Si_xSn_{1-x}$ solid solution. When two peaks were present, Otsu thresholding was applied to classify each pixel as either Mg_2Sn or Mg_2Si. Likewise, when three peaks were present, Otsu multi-thresholding was utilized for labelling the $Mg_2Si_xSn_{1-x}$, and Mg_2Si pixels.

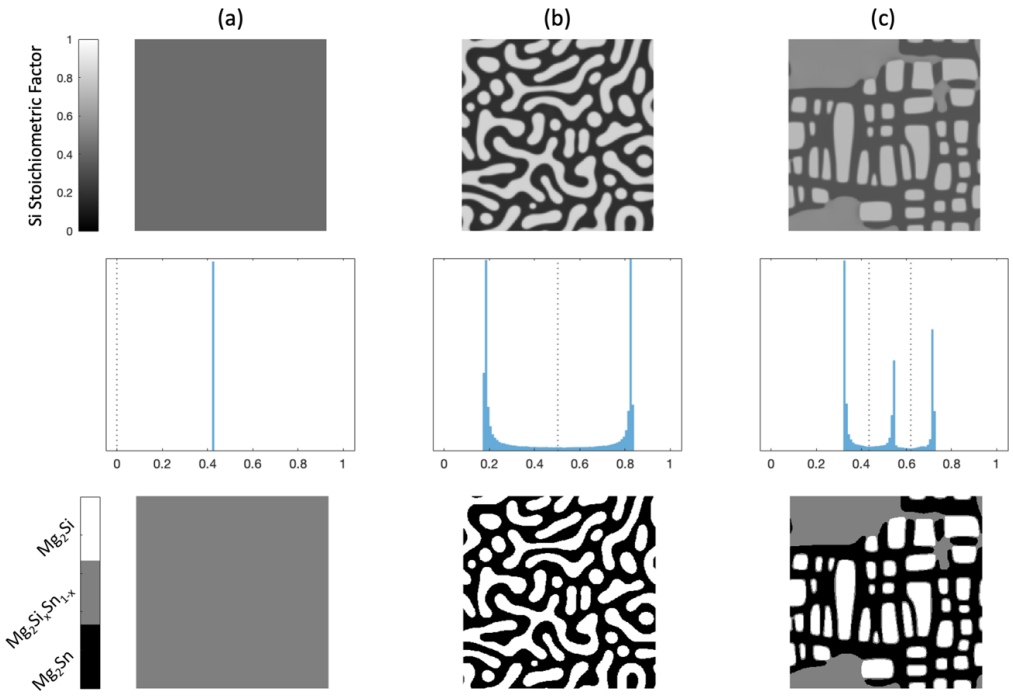

Figure 3. *Examples of segmentation of phase-field predicted microstructures: (a) homogeneous microstructure, (b) two-phase microstructure, and (c) three-phase microstructure. For each example, the top row shows the simulated microstructures, the middle row presents the histograms of the Si stoichiometric factor with dotted lines showing the thresholding limits, and the bottom row displays the segmented microstructures.*

It was determined that the OPMD dataset, at the selected evolution stage (15,000th time step), consisted of roughly 80% non-decomposed homogeneous samples, 14% fully decomposed two-phase samples, and 4% partially decomposed three-phase samples. The other 2% of the samples consisted primarily of physically unrealizable artifacts caused by simulation numerical

errors; these samples were removed from the dataset. We are interested in predicting the final microstructure of the phase field simulation; thus, the partially decomposed samples were also removed from the dataset.

3.3 Microstructure Quantification

The MKS framework was utilized for the quantification of the segmented two-phase microstructures. As already mentioned, the homogeneous (i.e., single phase) microstructures do not exhibit any morphological features. Note that each pixel in a segmented two-phase microstructure was assigned to one of the two possible local states: Mg_2Sn and Mg_2Si. As discussed in Section 2.1, only a single autocorrelation is required to capture all non-redundant 2-point spatial correlations for the two-phase microstructure [37]. In this work, the (Mg_2Sn-Mg_2Sn) autocorrelation, denoted simply as $f_{\boldsymbol{r}}$, is used. Fig. 4 shows examples of the computed autocorrelation maps for selected microstructures. The most prominent feature in each autocorrelation map is the central peak corresponding to $\boldsymbol{r} = \boldsymbol{0}$ (i.e., $f_{\boldsymbol{0}}$), which represents the volume fraction of the Mg_2Sn phase. Further morphological insights are found in the other autocorrelation peaks and valleys. For example, the autocorrelation patterns in examples 4(a) and 4(b) reflect the overall alignment of their respective morphological features; microstructures 4(a) and 4(b) are seen to have preferential Mg_2Sn (black phase) continuity along the diagonal direction and the horizontal and vertical directions, respectively. Meanwhile, the autocorrelation for microstructure 4(c) does not indicate any strong preferential directions in the alignment of the Mg_2Sn phase.

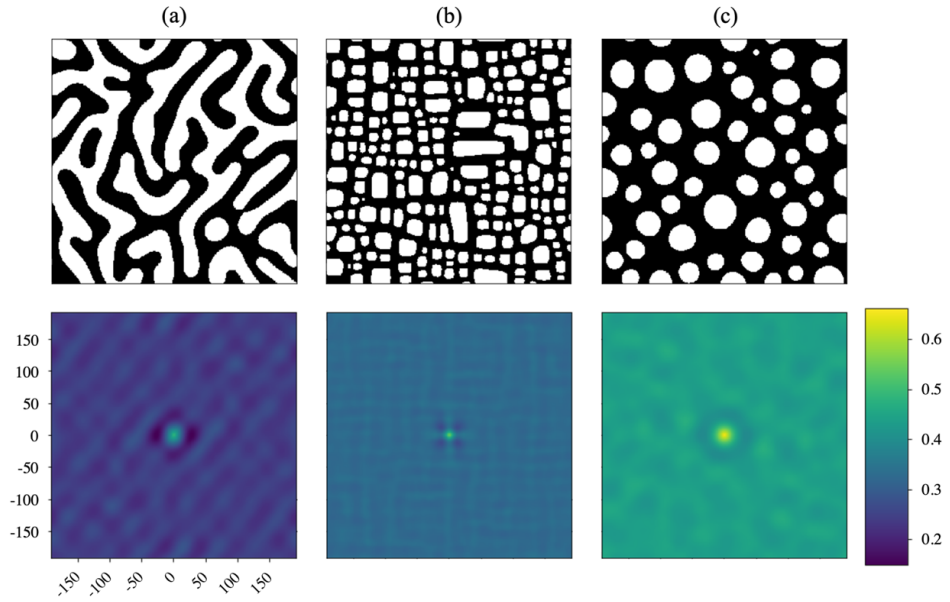

Figure 4. Mg_2Sn – Mg_2Sn autocorrelations (bottom) for example two-phase microstructures (top). Microstructure examples include (a) intertwined Mg_2Si and Mg_2Sn grains, (b) ordered cuboidal Mg_2Si nanoprecipitates, and (c) unordered spherical Mg_2Si nanoprecipitates. The value of the central peak in each map reflects the volume fraction of Mg_2Sn in the corresponding microstructure.

Each autocorrelation map in Fig. 4 contains 147,456 microstructure statistics (the same as the number of pixels in each simulated microstructure). We use PCA to produce a low-dimensional representation of this feature vector. The set of 1,365 two-phase microstructures was partitioned into a 85%/15% train/test split and PCA was performed on the autocorrelations in the training set (1,155 samples in total) using the scikit-learn package [49]. Over 99% of the dataset variance was captured in the first 5 PC scores and confirms the PCA analysis performed by Kunselman et al. [26] on a portion of the OPMD dataset. Fig. 5 shows reconstructed autocorrelations for an example microstructure using an increasing number of PC scores. It is seen that the reconstructed autocorrelation with 5 PC scores begins to closely resemble the original Mg_2Sn – Mg_2Sn autocorrelation and that further addition of PC scores does not greatly improve the accuracy of the reconstruction. Most samples in the dataset followed this trend and had diminishing returns on reconstruction accuracy past PC5.

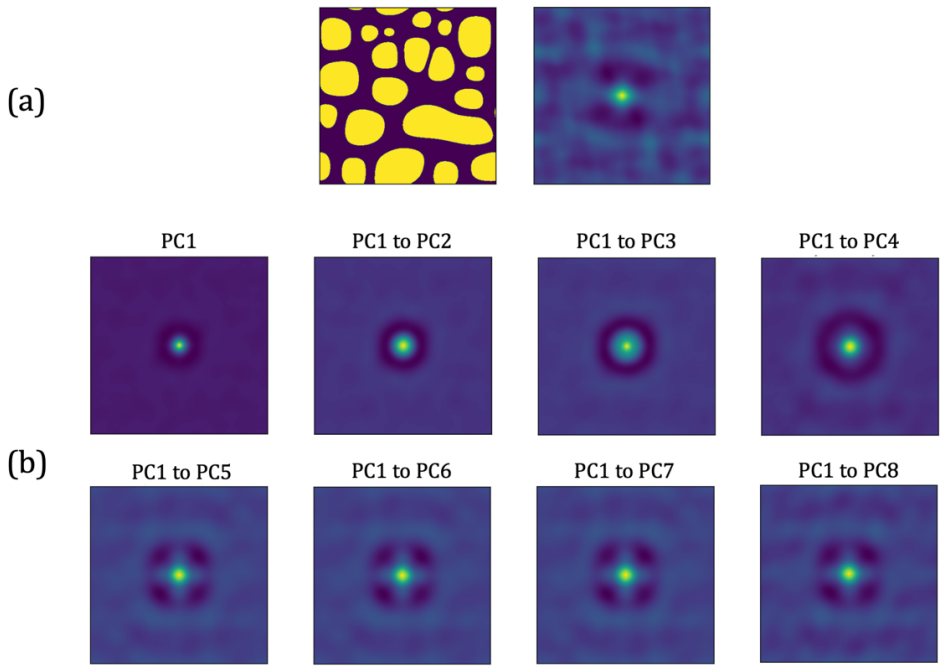

Figure 5. (a) *A two-phase microstructure and its Mg_2Sn – Mg_2Sn autocorrelation. (b) Reconstructed autocorrelations using different subsets of truncated PC representations. As an example, the PC1 to PC8 reconstruction signifies that the autocorrelation was reconstructed with only the first eight PC scores.*

The truncated PC representations produced a significant dimensionality reduction (from 147,456 spatial correlations to 5 PC scores) while still capturing most of the original information. A visualization of the distribution of the training set microstructures in the space of the first two PC scores is shown in Fig. 6. Each data point represents the autocorrelation of a single microstructure represented by its first two PC scores. Example microstructures are shown to demonstrate that the low-dimensional PC representations capture the salient morphological trends in the microstructure ensemble. Note that the figure only depicts the representations in the first two PC scores, while we have used five PC scores to represent the microstructures in this work. Consistent with prior applications [11,18], increases in PC1 were found to be highly correlated with increases in the Mg_2Sn volume fraction. Additionally, it was observed that PC2 correlated well with Mg_2Si precipitate size. Interestingly, microstructures exhibiting intertwined Mg_2Sn and Mg_2Si grains were found on the far-left side of the plot, while those with separated Mg_2Si precipitates were to the right. These results attest the ability of the MKS framework in providing

meaningful low-dimensional salient features for the quantitative representation of the microstructures.

While the protocols described above (microstructure → autocorrelation → PC scores) clearly capture large amounts of spatial information, they do represent a lossy filter. Higher-order spatial statistics (e.g., local grain shapes) are not included in the autocorrelation maps, and PCA truncation ensures that only the features capturing the highest dataset variance remain. As such, a predictive model built to learn the relationship between simulation input parameters and truncated PC scores would not be able to reproduce exact microstructure instantiations. Rather, the model would reproduce statistically similar microstructure autocorrelations which describe important spatial features (e.g., precipitate size and spacing).

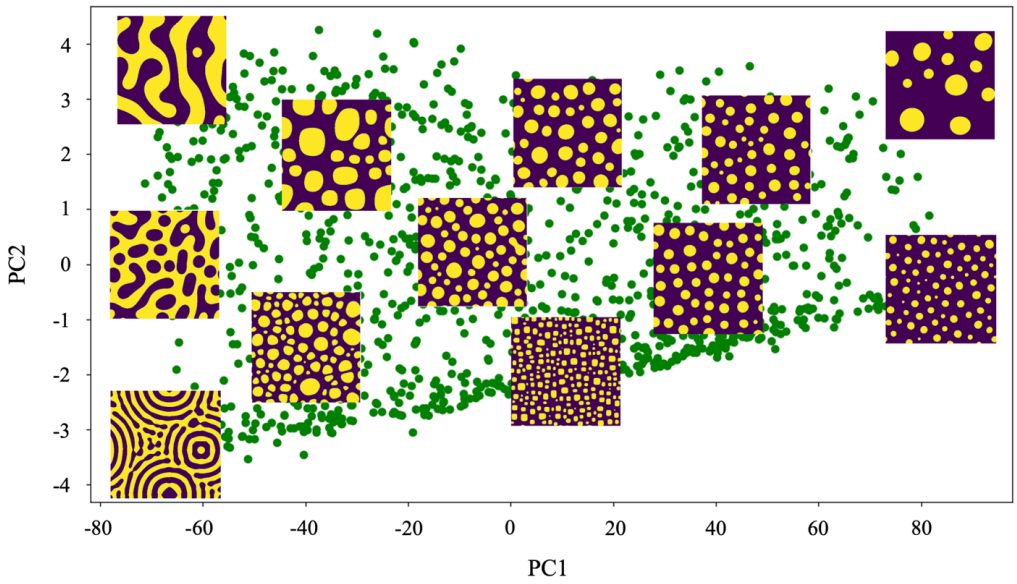

Figure 6. *Representation of the training set of two-phase microstructures (green dots) in the first two PC scores. Several two-phase microstructures (Mg_2Sn and Mg_2Si are shown purple and yellow, respectively) are overlayed directly above their location in PC space. While the first five PC scores are used for model building, much of the dataset morphological variation is captured in the first two components.*

3.4 Surrogate Model Building

The chained-ANN surrogate model built in this work consists of two MLPs (implemented with PyTorch [50]), as shown in Fig. 7. Each MLP was constructed with the same

train/validate/test split as described in Section 2.2 and the model hyperparameters for both MLPs were determined via heuristic refinement. The chained-ANN takes as input the phase-field simulation parameters (see Table 1), each of which has been mean centered and scaled to unit variance across all the samples in the training dataset to improve model training [51]. The transformed inputs were first fed to the MLP classifier, which distinguishes between sets of simulation parameters that result in homogeneous and heterogeneous microstructures. A predicted homogeneous microstructure is taken to be a solid solution of $Mg_2Si_xSn_{1-x}$, and no further analysis is performed. When the MLP classifier predicts a heterogeneous microstructure, the inputs are passed to the MLP regressor to predict the first five PC scores of microstructure autocorrelations. The PC scores are then used to reconstruct truncated autocorrelation maps (referred to as ML-predicted truncated autocorrelations) and corresponding microstructure instances.

As noted earlier, each ML-predicted truncated autocorrelation is expected to correspond to a relatively large set of microstructure instances, which exhibit similar two-point statistics [39,40]. At this point, it is important to recognize and understand the differences between microstructure instantiations and their representation using n-point spatial correlations; microstructure instantiations represent example microstructures that could have been sampled by an underlying process that defines the microstructure from its n-point spatial correlations. This underlying stochastic process is completely described by the full set of n-point spatial correlations, of which the autocorrelation maps constitute a first-order approximation. As such, the framework of spatial correlations offers a more accurate and meaningful representation of the microstructure, while any single instantiation should only be treated as a possible sampled microstructure. For this reason, spatial correlations of the microstructure should be the preferred choice for the representation of the microstructure in PSP models used to drive materials development efforts [6].

In this work, we present example instantiations corresponding to the ML-predicted truncated autocorrelation maps, produced via an efficient Gaussian Random Field (GRF) generator [40]. Due to the stochastic nature of this procedure, these are unlikely to match the phase-field predicted microstructures. Nevertheless, it is instructive to examine what spatial features of the phase-field predicted microstructures are indeed captured in the first five PC scores used in the ML-predicted truncated autocorrelation maps.

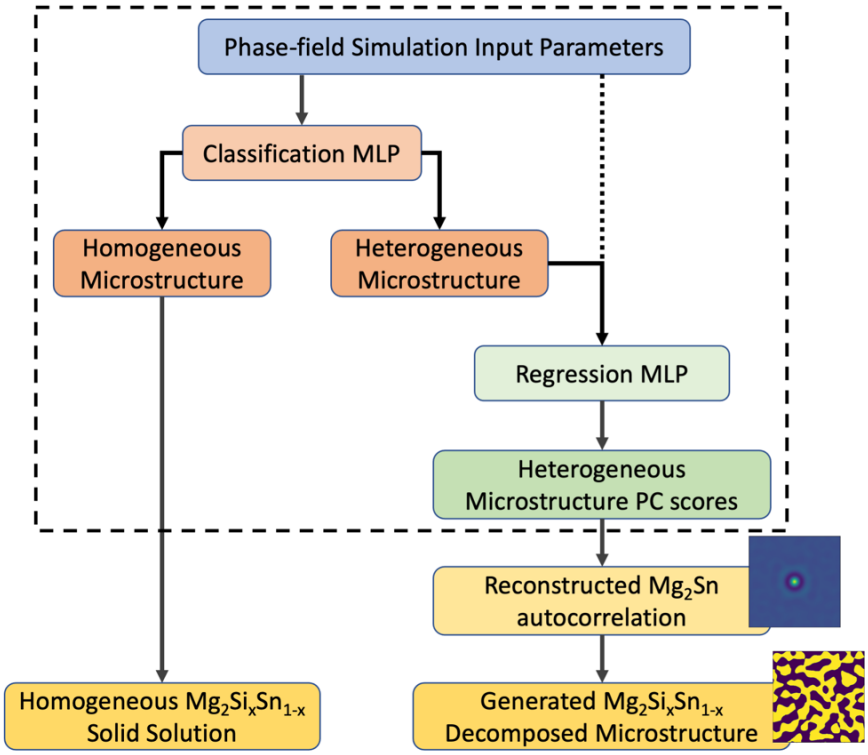

Figure 7. A schematic showing the chained-ANN process-structure surrogate model developed in this work. The chained-ANN model (inside the dashed box) contains both the classification model (orange) and the regression model (green). Microstructure information extracted from the model outputs are shown in yellow.

3.4.1 Classification MLP

The ideal loss function was found to be Binary Cross Entropy (BCE) loss [52], common for models with binary outputs (in our case, zero and one correspond to homogeneous and heterogeneous microstructures, respectively). The BCE loss function was minimized with the AdamW gradient descent optimizer [53] with an initial learning rate of 0.001 and a weight decay of 1e-2. The ideal model size, corresponding to the smallest number of parameters (see Eq. 2) with the highest accuracy, was found to be a single hidden layer with five nodes, containing only 101 fittable parameters. Larger models (either in depth or layer width) tended to aggressively overfit and produce worse overall validation set performance. The ReLU activation function [43] was used on the hidden layer, and the sigmoid activation function [43] applied to the output layer to ensure that the predicted values lie between zero and one. The network learning rate was governed

via the 1cycle learning rate policy [54], which briefly increased the learning rate before decreasing to a rate much lower than the originally prescribed value of 0.001. The 1cycle learning rate helped the model converge quickly (see Fig. 8), performed well over a wide range of total training epochs, and improved model stability (i.e., more consistent results across multiple training attempts). The ideal ratio of model performance to training time was achieved with fifty training epochs.

After model training was complete, the predicted values were rounded to the nearest integer (0.5 thresholding) and the classification accuracy was quantified using the precision, recall, and $f1$-score (see Eqs. 3-5). Because the rounding threshold used to binarize the classifier output is an essential model parameter, a sensitivity analysis was performed to determine the ideal thresholding value. The sensitivity analysis showed that deviation from 0.5 resulted in slightly worse $f1$-score measurements; however, this decrease in model performance was minor over the range of 0.2-0.8, likely due to the inherent symmetric bias from the sigmoid activation function toward zero and one.

As the classification MLP is the first predictive network in the chained-ANN, the classification error is propagated through to the regression network and into the autocorrelation reconstructions; thus, it is particularly important that the classification model be as accurate as possible before chaining it with the MLP regressor. Because the OPMD dataset consists of primarily homogenous microstructures, the model accuracy is naturally weighted towards the classification accuracy of homogeneous samples. This weighted prediction accuracy can be seen clearly in Table 2 where the training and testing dataset classifications are presented categorically in a confusion matrix. Also evident in Table 2 is the ratio of homogeneous and heterogeneous samples contained in each dataset partition. Roughly 15% of the samples in both the training and testing datasets are from decomposed simulations. The precision, recall, and $f1$-score for both homogeneous and heterogeneous classifications are summarized in Table 3 for the training and test dataset partitions. The homogeneous classification is highly accurate with a test set $f1$-score of 0.996. Similarly, the heterogeneous classification is highly accurate with a test set $f1$-score of 0.976.

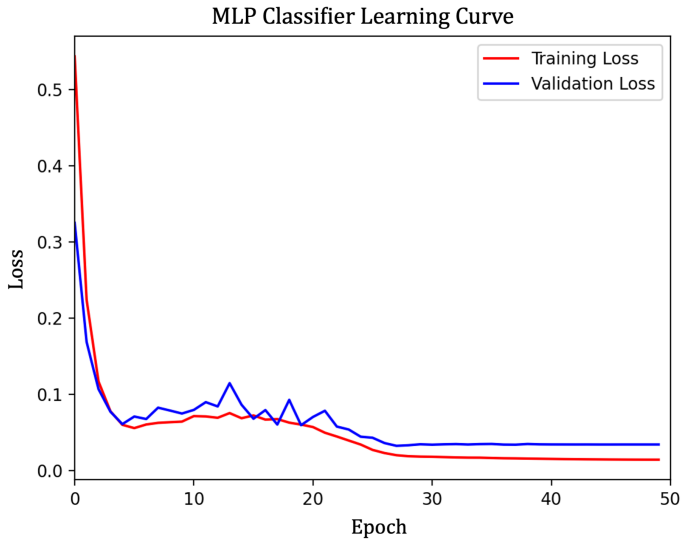

Figure 8. Learning curves containing training and validation loss for the classification MLP.

Table 2. Classification confusion matrix for both the training and testing dataset partitions.

	Training Dataset	
	Predicted Homogeneous	Predicted Heterogeneous
True Homogeneous	6909	23
True Heterogeneous	34	1121

	Testing Dataset	
	Predicted Homogeneous	Predicted Heterogeneous
True Homogeneous	1216	2
True Heterogeneous	8	202

Table 3. The average precision, recall, and f1-score for the constructed classification model for both homogeneous and heterogeneous (see Eqs. 3-5).

	Homogeneous Classification		
	Precision	Recall	*f*1-Score
Train Set	0.995	0.997	0.996
Test Set	0.993	0.998	0.996

	Heterogeneous Classification		
	Precision	Recall	*f*1-Score
Train Set	0.980	0.971	0.975
Test Set	0.990	0.962	0.976

3.4.2 Regression MLP

The regression model outputs (i.e., the first five PC scores) were scaled to lie between the values of -1.0 and 1.0 to facilitate model training [51]. The Huber loss function [52], which is less sensitive to outliers, was found to give the best results. The loss function was minimized with the AdamW optimizer with a learning rate of 5e-4 and a weight decay of 5e-3. The GELU activation function and batch normalization were utilized in each hidden layer [43,55]. The number and size of the hidden layers were varied to determine the ideal model architecture. It was found that three hidden layers with twelve nodes each (corresponding to 641 fittable parameters) produced the ideal regression model for our case study. The network learning rate, similar to the classification MLP, was governed by the 1cycle learning rate policy and helped the regression model converge quickly to a local optimum (see Fig. 9). The ideal number of training epochs, resulting in the lowest validation loss, was determined to be two hundred.

After model training was completed, the predicted scaled PC scores were converted back to their original domains and model prediction accuracy was calculated using Eq. 6. The $nMAE$ (see Eq. 6) for each PC score is reported in Table 4. The $nMAE$ increases for higher PC scores, as expected, because higher PC scores (associated with lower variance PC basis) tend to capture more noise [56]. This is also seen in the parity plots in Fig. 10 where the higher PC scores become less accurate in both training and test dataset partitions and, by PC5, the predictions have poor correlation to the true values. While the higher PC score predictions are less accurate, they also hold less importance, as the PC scores are organized by their ability to capture the dataset variance. Autocorrelation reconstructions from the predicted PC scores will depend mostly on the more accurate (PC1-PC3) scores, while the less accurate scores will have a much smaller impact.

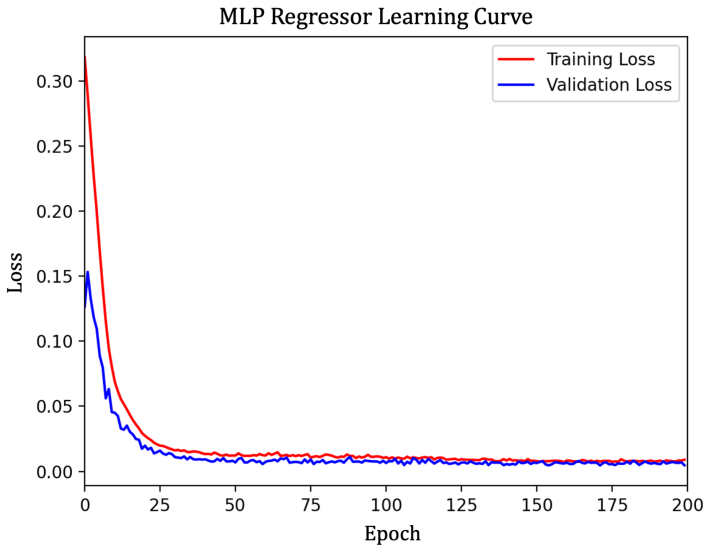

Figure 9. Learning curves containing training and validation loss for the regression MLP.

Table 4. Reported $nMAE$ (see Eq. 6) and standard deviations for the predicted PC scores for the train and test dataset partitions.

	$nMAE$				
	PC1	PC2	PC3	PC4	PC5
Train Set	0.086 ± 0.111	0.175 ± 0.165	0.310 ± 0.292	0.483 ± 0.437	0.790 ± 0.875
Test Set	0.108 ± 0.227	0.213 ± 0.204	0.351 ± 0.292	0.561 ± 0.509	0.996 ± 1.393

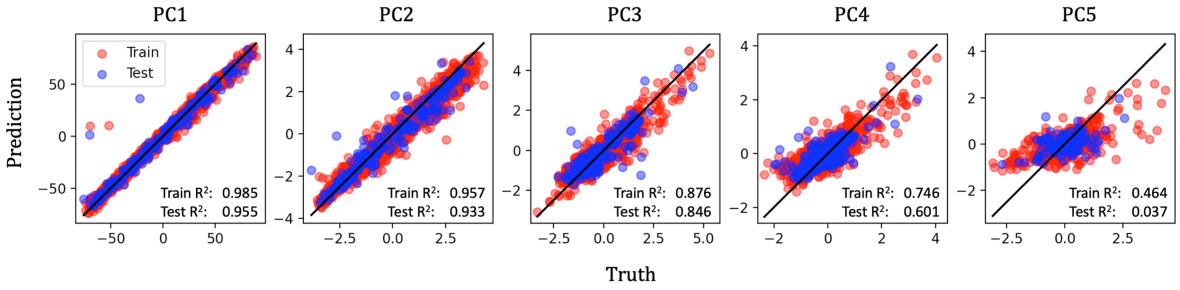

Figure 10. Parity plots of the MLP Regressor PC score predictions for the train (red) and test (blue) dataset partitions. The R^2 values are given for each parity plot.

Final model accuracy was determined by quantifying the difference between the ML-predicted truncated autocorrelations ($\tilde{\boldsymbol{y}}$) and the autocorrelations of the corresponding phase-field predicted microstructures ($\boldsymbol{y}$) as

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \tilde{y}_i)^2}{N}} \tag{7}$$

where N is the total number of autocorrelation features. The test set $RMSE$ is presented in Fig. 11 as a frequency plot along with three example microstructures specifically selected to demonstrate the range of model accuracy. Samples (a) and (b) correspond to minimum and near-mean $RMSE$. It is seen that the ML-predicted truncated autocorrelations for these samples are qualitatively very similar to the truncated autocorrelations of the corresponding phase-field predicted microstructures. Sample (c) represents one of the test set outliers with high reconstruction error, and clearly does not adequately represent the corresponding truncated autocorrelation of the phase-field predicted microstructure. The relatively high error associated with this, and other outliers is attributed to being substantially distant from the training data points (in the PC space). Specifically, for sample (c), the PC3 value was substantially outside of the range of PC3 values in the training set. Similar observations were made for the other test set outliers exhibiting higher RMSE values. As with all applications of ML, the goal is to construct a model that is well generalizable to unseen data; however, ML models typically excel at data interpolation and normally perform worse at extrapolating outside of the range of training set. As such, the reconstruction outliers, which were determined to exist outside of the training set PC space, can be attributed to the MLP regressor's inaptitude for extrapolation. This points to the need for improved design of training sets for ML models to promote more generalizable models. Notwithstanding, all three ML-predicted truncated autocorrelations adequately captured Mg_2Sn volume fraction (i.e., $f_{\mathbf{0}}$). The absolute volume fraction errors for samples (a), (b) and (c) were 0.25, 1.17, and 4.11 wt.% Mg_2Sn, respectively. The mean absolute error in volume fraction predictions for the whole test set was 1.06 wt.% Mg_2Sn.

While the differences between the ML-predicted truncated autocorrelations and the truncated autocorrelations of the phase-field predicted microstructures for samples (a) and (b) are minimal, we see that the ML-predicted truncated autocorrelations lack the long-range order present in their respective original autocorrelations. This is a consequence of PCA truncation. To visualize

the deficiencies of the ML-predicted truncated autocorrelations, the recently developed [40] GRF microstructure generator was employed on each of the ML-predicted truncated autocorrelation maps in Fig. 11. While only a single generated microstructure instance for each sample is shown, we produced a multitude of microstructures for each autocorrelation map. It was observed that the generated instantiations captured certain aspects of each microstructure very well (such as the phase volume fractions and the general size distribution of precipitates). Notably, Mg_2Sn phase volume fraction and Mg_2Si precipitate size are the two features which were readily expressed in PC1 and PC2, respectively (See Fig. 6). However, the generated microstructures do not capture well the shapes of the individual precipitates. This is because local shape descriptors are generally found in higher-order spatial correlations [40]. Furthermore, the GRF generator used in this study is known to be implicitly biased toward the generation of microstructures with semi-continuous features due to its inherent Gaussian assumptions. To address these issues, the overall strategy described here can be extended in future studies to include higher-order spatial correlations.

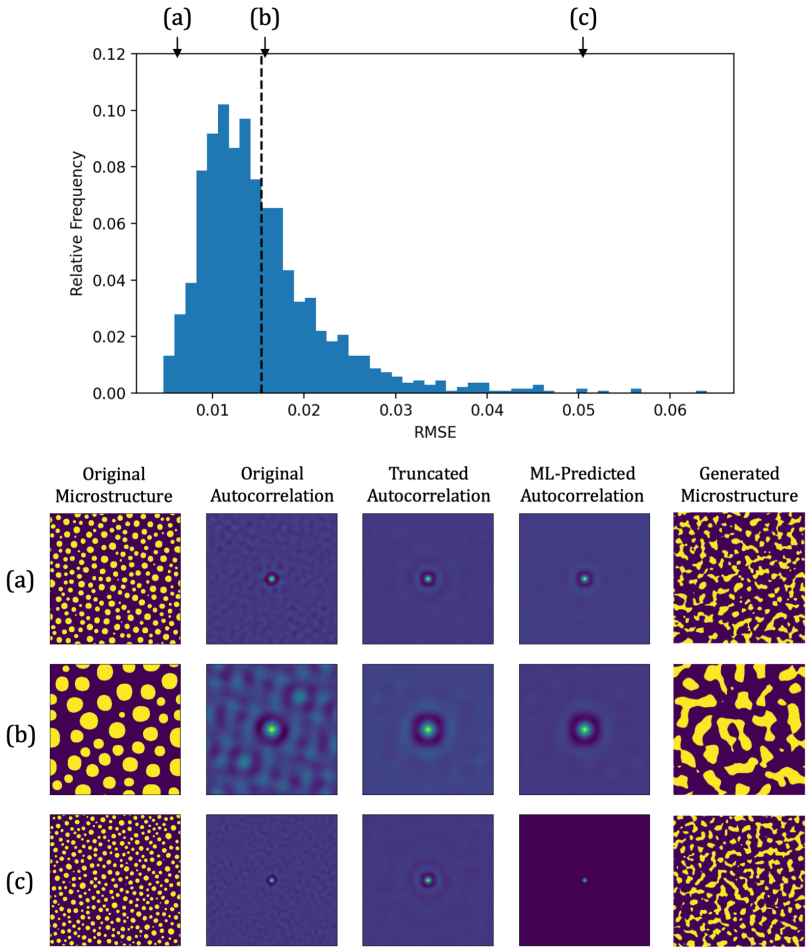

Figure 11. RMSE probability density for the test set ML-predicted truncated autocorrelations (top). The original, truncated, and predicted (reconstructed from predicted PC scores) autocorrelations for samples with (a) minimum reconstruction error, (b) near-mean reconstruction error, and (c) maximum reconstruction error (bottom). The example GRF generated microstructures were produced using the ML-predicted truncated autocorrelations.

3.4.3 Chained-ANN Surrogate Model

The trained classification and regression MLPs were chained as depicted in Fig. 7. The chained-ANN process-structure surrogate model accuracy was measured over the test dataset partition. If the classification model were perfect, meaning all samples in the test dataset partition were correctly classified, then the chained-ANN surrogate model would exhibit the same regression accuracy as the individual regression MLP. However, in the chained-ANN, the classification error is propagated to the regressor and contributes to the final error of the ML-predicted truncated autocorrelations. The classification model incorrectly labeled 0.62% of the homogeneous samples as heterogeneous and 5.94% of the heterogeneous samples as homogeneous. Nearly one in twenty heterogeneous microstructure producing simulation inputs were labeled as homogeneous and weren't included in the subsequent PC score predictions. Equivalently, nearly one in two-hundred homogeneous microstructure producing simulation inputs were labeled as heterogeneous and their PC scores were predicted by the regressor MLP.

The false heterogeneous samples become clearly evident in some of the PC score prediction parity plots shown in Fig. 12 (circled in red where identifiable as a single group). Because the regression MLP was trained solely on heterogeneous samples, the false heterogeneous classifications contribute significantly to the overall model error. The $nMAE$ values provided in Table 4 reflect the loss of accuracy due to the low classification error from the false heterogeneous samples as well as the four omitted true heterogeneous samples. Since PC4 and PC5 capture more of the autocorrelation noise, the contribution from the classification error is fairly minimal to their predictions (see Fig. 12 and Table 5). As with the reconstruction outliers discussed in Section 3.4.2, we see that the regression MLP is attempting to assign interpolated PC scores to the homogeneous samples, which were clearly and purposefully outside of the regression MLP training set, rather than extrapolating to their true values. This is expected and is a result of the

chained-ANN design decisions made in this work. As classification error decreases, the overall chained-ANN prediction error will converge to the same error seen with the regression MLP in Table 4.

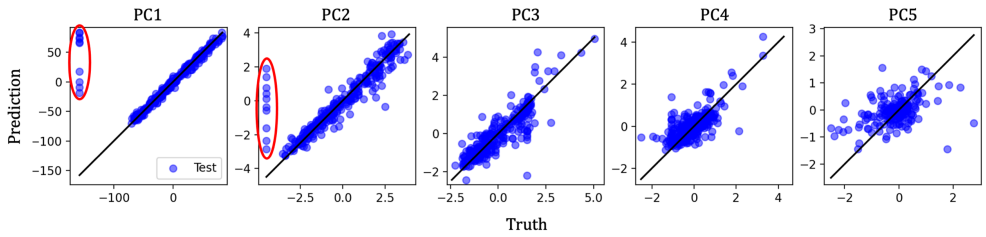

Figure 12. Parity plots of the chained-ANN PS surrogate model PC score predictions for the test dataset partitions. Homogeneous samples which were incorrectly classified as heterogeneous are circled in red where visible.

Table 5. Reported nMAE (see Eq. 5) and standard deviations for the predicted PC scores of the test dataset partition as measured with the MLP Regressor and the chained-ANN PS surrogate model

	$nMAE$				
	PC1	PC2	PC3	PC4	PC5
MLP Regressor	0.108 ± 0.227	0.213 ± 0.204	0.351 ± 0.292	0.561 ± 0.509	0.996 ± 1.393
Chained-ANN	0.266 ± 0.965	0.273 ± 0.483	0.389 ± 0.418	0.657 ± 0.642	0.907 ± 1.188

Although there is aggregated error from the classification model in the chained-ANN output, the surrogate model obtained using this framework remains a promising opportunity for the acceleration of the computationally expensive phase-field simulations. The chained-ANN surrogate model can determine heterogeneity and predict a final Mg_2Sn- Mg_2Sn autocorrelation in only 65 milliseconds (3.5 GHz Dual-Core Intel Core i7), while the phase-field simulation takes roughly 2.5 hours (2.4 GHz Quad-Core Intel Core i5) to produce a single simulated microstructure instantiation (representing a computational efficiency increase of five orders of magnitude). The benefit of the autocorrelation surrogate model output, as noted previously, is that a host of microstructures can be generated which are all statistically similar to the predicted microstructure statistics rather than being limited to a single microstructure instantiation. Furthermore, the newly developed surrogate model allows for quick prediction of phase separation based on simulation

input parameters. The OPMD dataset consists of roughly 80% non-decomposed microstructures; implying that roughly four in five sets of simulation inputs could be ignored when searching for heterogeneous structures. The gained computational efficiency along with the ability to quickly sift through homogeneous and heterogeneous producing simulation parameters allows for exceptional opportunities for high-throughput explorations of the $Mg_2Si_xSn_{1-x}$ microstructure space. For example, with the goal of developing more efficient thermoelectric $Mg_2Si_xSn_{1-x}$ microstructures, the surrogate model can be used to efficiently explore a vast and high-dimensional parameter space to identify which regions will produce heterogeneous microstructures. Further work could be done to create an inverse mapping of the heterogeneous microstructure space allowing for ideal process parameters to be identified before running the computationally expensive phase-field simulations to obtain more detailed microstructure predictions.

4. Conclusions

A novel approach utilizing the MKS framework with a chained-ANN workflow is presented for learning a complex, highly nonlinear, and high-dimensional process-structure mapping from phase-field simulation data. This work offers a distinct new avenue for learning increasingly complex process-structure relationships, while using very few model fitting parameters. For the selected case study involving phase-field simulations of spinodal decomposition in the $Mg_2Si_xSn_{1-x}$ material system, the novel approach presented in this work required only 742 model fitting parameters to map an 18-dimensional process space to its corresponding microstructure space. The number of model parameters is not only significantly smaller than the number of training samples (8,087 for the present study), but is several orders of magnitude smaller than those for surrogate models reported in current literature for similar datasets.

Specifically, the benefits of the proposed model building strategy have been demonstrated on a publicly available dataset of $Mg_2Si_xSn_{1-x}$ spinodal decomposition phase-field simulations which use eighteen unique input parameters to simulate microstructures from a highly diverse microstructure space. The chained-ANN surrogate model differentiated between homogeneous and heterogeneous producing simulation inputs with high accuracy and predicted two-phase spatial autocorrelations with close agreement to the training data. We note here that the use of PCA

truncation removed higher-order features from the autocorrelations used in model training and limited the autocorrelation reconstruction accuracy considerably. Additional dimensionality reduction techniques should be explored in future work to overcome this barrier. While the model building strategies used in this case study are specific to the OPMD spinodal decomposition dataset, the machine learning framework described herein opens new research avenues for an unprecedented capture of complex process-structure mappings over very large high-dimensional input and output spaces in a broad range of material systems.

Acknowledgements

The authors would like to thank Berkay Yucel and Andreas Robertson for their helpful discussions throughout this work. This work was supported by Sandia LDRD Contract Agreement #2177713 and NSF Graduate Research Fellowship DGE-1650044. VA and RA would like to acknowledge NSF through Grants No. NSF-CDSE-2001333 and NSF-DMR-1905325, as well as the High-Performance Research Computing Facility at their institution. Any opinion, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of Sandia National Laboratories or the National Science Foundation. The implementation and training code for each MLP and the full chained-ANN are available under an open-source license on GitHub [57].

References

[1] P. Voorhees, G. Spanos, others, Modeling across scales: a roadmapping study for connecting materials models and simulations across length and time scales, TMS, Warrendale, PA. 14 (2015).
[2] Materials Genome Initiative for Global Competitiveness, 2011. www.ostp.gov/nstc.
[3] P. Cummings, S. Glotzer, Inventing a New America Through Discovery and Innovation in Science, Engineering and Medicine: A Vision for Research and Development in Simulation-Based Engineering and Science in the Next Decade, 2010. https://www.nsf.gov/mps/ResearchDirectionsWorkshop2010/RWD-color-FINAL-usletter_2010-07-16.pdf (accessed December 26, 2021).
[4] S. Glotzer, S. Kim, P. Cummings, A. Deshmukh, M. Head, G. Karniadakis, L. Petzold, C. Sagui, M. Shinozuka, WTEC Panel Report on International Assessment Of Research And Development In Simulation-Based Engineering And Science, (2009).
[5] N.R. Council, others, Integrated computational materials engineering: a transformational discipline for improved competitiveness and national security, National Academies Press, 2008.

[6] S.R. Kalidindi, Hierarchical materials informatics: novel analytics for materials data, Elsevier, 2015.
[7] D.T. Fullwood, S.R. Niezgoda, B.L. Adams, S.R. Kalidindi, Microstructure Sensitive Design for Performance Optimization, Progress in Materials Science. 55 (2010) 477–562. https://doi.org/10.1016/J.PMATSCI.2009.08.002.
[8] D.B. Brough, D. Wheeler, J.A. Warren, S.R. Kalidindi, Microstructure-Based Knowledge Systems for Capturing Process-Structure Evolution Linkages, Current Opinion in Solid State and Materials Science. 21 (2017) 129–140. https://doi.org/10.1016/J.COSSMS.2016.05.002.
[9] D. Sarkar, B.S. Reddy, B. Basu, Implementing statistical modeling approach towards development of ultrafine grained bioceramics: Case of ZrO2-toughened Al2O3, Journal of the American Ceramic Society. 101 (2018) 1333–1343.
[10] 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, Journal of Manufacturing Science and Engineering. 139 (2017).
[11] Y.C. Yabansu, P. Steinmetz, J. Hötzer, S.R. Kalidindi, B. Nestler, Extraction of Reduced-Order Process-Structure Linkages from Phase-Field Simulations, Acta Materialia. 124 (2017) 182–194. https://doi.org/10.1016/j.actamat.2016.10.071.
[12] A. Cecen, H. Dai, Y.C. Yabansu, S.R. Kalidindi, L. Song, Material structure-property linkages using three-dimensional convolutional neural networks, Acta Materialia. 146 (2018) 76–84.
[13] E. Popova, T.M. Rodgers, X. Gong, A. Cecen, J.D. Madison, S.R. Kalidindi, Process-Structure Linkages Using a Data Science Approach: Application to Simulated Additive Manufacturing Data, Integrating Materials and Manufacturing Innovation. 6 (2017) 54–68. https://doi.org/10.1007/s40192-017-0088-1.
[14] I. Hassinger, X. Li, H. Zhao, H. Xu, Y. Huang, A. Prasad, L. Schadler, W. Chen, L. Catherine Brinson, Toward the Development of a Quantitative Tool for Predicting Dispersion of Nanocomposites under Non-equilibrium Processing Conditions, Journal of Materials Science. 51 (2016) 4238–4249. https://doi.org/10.1007/s10853-015-9698-1.
[15] S. Gorgannejad, M. Reisi Gahrooei, K. Paynabar, R.W. Neu, Quantitative prediction of the aged state of Ni-base superalloys using PCA and tensor regression, Acta Materialia. 165 (2019) 259–269. https://doi.org/10.1016/j.actamat.2018.11.047.
[16] D.B. Brough, D. Wheeler, S.R. Kalidindi, Materials knowledge systems in python—a data science framework for accelerated development of hierarchical materials, Integr Mater Manuf Innov. 6 (2017) 36–53.
[17] S. Hashemi, S.R. Kalidindi, A Machine Learning Framework for the Temporal Evolution of Microstructure During Static Recrystallization of Polycrystalline Materials Simulated by Cellular Automaton, Computational Materials Science. 188 (2021). https://doi.org/10.1016/j.commatsci.2020.110132.
[18] S. Hashemi, B. Ganapathysubramanian, S. Casey, J. Su, S.R. Kalidindi, Feature engineering for microstructure-property mapping in organic photovoltaics, ArXiv Preprint ArXiv:2111.01897. (2021).
[19] P.R. Kaundinya, K. Choudhary, S.R. Kalidindi, Prediction of the electron density of states for crystalline compounds with Atomistic Line Graph Neural Networks (ALIGNN), ArXiv Preprint ArXiv:2201.08348. (2022).

[20] S. Pfeifer, O. Wodo, B. Ganapathysubramanian, An optimization approach to identify processing pathways for achieving tailored thin film morphologies, Computational Materials Science. 143 (2018) 486–496.
[21] S. Timmermann, V. Starostin, A. Girelli, A. Ragulskaya, H. Rahmann, M. Reiser, N. Begam, L. Randolph, M. Sprung, F. Westermeier, others, Automated matching of two-time X-ray photon correlation maps from phase-separating proteins with Cahn–Hilliard-type simulations using auto-encoder networks, Journal of Applied Crystallography. 55 (2022).
[22] G.H. Teichert, A.R. Natarajan, A. der Ven, K. Garikipati, Machine learning materials physics: Integrable deep neural networks enable scale bridging by learning free energy functions, Computer Methods in Applied Mechanics and Engineering. 353 (2019) 201–216.
[23] G.H. Teichert, K. Garikipati, Machine learning materials physics: Surrogate optimization and multi-fidelity algorithms predict precipitate morphology in an alternative to phase field dynamics, Computer Methods in Applied Mechanics and Engineering. 344 (2019) 666–693.
[24] X. Zhang, K. Garikipati, Machine learning materials physics: Multi-resolution neural networks learn the free energy and nonlinear elastic response of evolving microstructures, Computer Methods in Applied Mechanics and Engineering. 372 (2020) 113362.
[25] M.F. Badawy, M.A. Msekh, K.M. Hamdia, M.K. Steiner, T. Lahmer, T. Rabczuk, Hybrid nonlinear surrogate models for fracture behavior of polymeric nanocomposites, Probabilistic Engineering Mechanics. 50 (2017) 64–75.
[26] C. Kunselman, V. Attari, L. McClenny, U. Braga-Neto, R. Arroyave, Semi-supervised learning approaches to class assignment in ambiguous microstructures, Acta Materialia. 188 (2020) 49–62.
[27] C. Kunselman, S. Sheikh, M. Mikkelsen, V. Attari, R. Arróyave, Microstructure classification in the unsupervised context, Acta Materialia. 223 (2022) 117434.
[28] A.A. Kazemzadeh Farizhandi, M. Mamivand, Processing Time, Temperature, and Initial Chemical Composition Prediction from Microstructure by Deep Neural Network for Multiple Inputs and Fused Data, Temperature, and Initial Chemical Composition Prediction from Microstructure by Deep Neural Network for Multiple Inputs and Fused Data. (n.d.).
[29] E. Herman, J.A. Stewart, R. Dingreville, A data-driven surrogate model to rapidly predict microstructure morphology during physical vapor deposition, Applied Mathematical Modelling. 88 (2020) 589–603. https://doi.org/10.1016/J.APM.2020.06.046.
[30] C. Hu, S. Martin, R. Dingreville, Accelerating phase-field predictions via recurrent neural networks learning the microstructure evolution in latent space, Computer Methods in Applied Mechanics and Engineering. 397 (2022) 115128.
[31] D. Montes de Oca Zapiain, J.A. Stewart, R. Dingreville, Accelerating phase-field-based microstructure evolution predictions via surrogate models trained by machine learning methods, Npj Computational Materials. 7 (2021). https://doi.org/10.1038/s41524-020-00471-8.
[32] V. Attari, D.J. Sauceda, Open Phase-field Microstructure Database (OPMD), (2019). http://microstructures.net/.

[33] V. Attari, P. Honarmandi, T. Duong, D.J. Sauceda, D. Allaire, R. Arroyave, Uncertainty Propagation in a Multiscale CALPHAD-Reinforced Elastochemical Phase-Field Model, Acta Materialia. 183 (2020) 452–470. https://doi.org/10.1016/j.actamat.2019.11.031.
[34] L. McClenny, M. Haile, V. Attari, B. Sadler, U. Braga-Neto, R. Arroyave, Deep Multimodal Transfer-Learned Regression in Data-Poor Domains, ArXiv Preprint ArXiv:2006.09310. (2020).
[35] B.L. Adams, X. Gao, S.R. Kalidindi, Finite Approximations to the Second-Order Properties Closure in Single Phase Polycrystals, Acta Materialia. 53 (2005) 3563–3577. https://doi.org/10.1016/J.ACTAMAT.2005.03.052.
[36] S.R. Niezgoda, A.K. Kanjarla, S.R. Kalidindi, Novel Microstructure Quantification Framework for Databasing, Visualization, and Analysis of Microstructure Data, Integrating Materials and Manufacturing Innovation. 2 (2013) 54–80. https://doi.org/10.1186/2193-9772-2-3.
[37] S.R. Niezgoda, D.T. Fullwood, S.R. Kalidindi, Delineation of the Space of 2-Point Correlations in a Composite Material System, Acta Materialia. 56 (2008) 5285–5292. https://doi.org/10.1016/J.ACTAMAT.2008.07.005.
[38] A.P. Generale, S.R. Kalidindi, Reduced-order models for microstructure-sensitive effective thermal conductivity of woven ceramic matrix composites with residual porosity, Composite Structures. 274 (2021) 114399.
[39] D.T. Fullwood, S.R. Niezgoda, S.R. Kalidindi, Microstructure reconstructions from 2-point statistics using phase-recovery algorithms, Acta Materialia. 56 (2008) 942–948. https://doi.org/10.1016/j.actamat.2007.10.044.
[40] A.E. Robertson, S.R. Kalidindi, Efficient Generation of Anisotropic N-field Microstructures from 2-Point Statistics using Multi-output Gaussian Random Fields, Acta Materialia. (2022) 117927.
[41] H. Izadi, M. Baniassadi, A. Hasanabadi, B. Mehrgini, H. Memarian, H. Soltanian-Zadeh, K. Abrinia, Application of full set of two point correlation functions from a pair of 2D cut sections for 3D porous media reconstruction, Journal of Petroleum Science and Engineering. 149 (2017) 789–800.
[42] A. Hasanabadi, M. Baniassadi, K. Abrinia, M. Safdari, H. Garmestani, 3D microstructural reconstruction of heterogeneous materials from 2D cross sections: A modified phase-recovery algorithm, Computational Materials Science. 111 (2016) 107–115. https://doi.org/10.1016/j.commatsci.2015.09.015.
[43] S.R. Dubey, S.K. Singh, B.B. Chaudhuri, A Comprehensive Survey and Performance Analysis of Activation Functions in Deep Learning, ArXiv Preprint ArXiv:2109.14545. (2021).
[44] R. Hecht-Nielsen, Theory of the backpropagation neural network, in: Neural Networks for Perception, Elsevier, 1992: pp. 65–93.
[45] 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 2 Si–Mg 2 Sn solid solutions, Journal of Materials Chemistry A. 6 (2018) 17559–17570.
[46] S. Wang, N. Mingo, Improved Thermoelectric Properties of Mg2SixGeySn1−X−Y Nanoparticle-In-Alloy Materials, Applied Physics Letters. 94 (2009). https://doi.org/10.1063/1.3139785.

[47] Q. Zhang, J. He, T.J. Zhu, S.N. Zhang, X.B. Zhao, T.M. Tritt, High Figures of Merit and Natural Nanostructures in Mg2Si0.4Sn0.6 Based Thermoelectric Materials, Applied Physics Letters. 93 (2008). https://doi.org/10.1063/1.2981516.
[48] N. Otsu, A threshold selection method from gray-level histograms, IEEE Trans Syst Man Cybern. 9 (1979) 62–66.
[49] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, others, Scikit-learn: Machine learning in Python, The Journal of Machine Learning Research. 12 (2011) 2825–2830.
[50] A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, others, Pytorch: An imperative style, high-performance deep learning library, Adv Neural Inf Process Syst. 32 (2019).
[51] Y.A. LeCun, L. Bottou, G.B. Orr, K.-R. Müller, Efficient backprop, in: Neural Networks: Tricks of the Trade, Springer, 2012: pp. 9–48.
[52] Q. Wang, Y. Ma, K. Zhao, Y. Tian, A comprehensive survey of loss functions in machine learning, Annals of Data Science. 9 (2022) 187–212.
[53] I. Loshchilov, F. Hutter, Fixing weight decay regularization in adam, (2018).
[54] L.N. Smith, N. Topin, Super-convergence: Very fast training of neural networks using large learning rates, in: Artificial Intelligence and Machine Learning for Multi-Domain Operations Applications, 2019: pp. 369–386.
[55] S. Ioffe, C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, in: International Conference on Machine Learning, 2015: pp. 448–456.
[56] A.E. Robertson, S.R. Kalidindi, Digital Representation and Quantification of Discrete Dislocation Structures, JOM. 73 (2021) 2143–2158.
[57] G. Harrington, C. Kelly, Chained-ANN-Phase-field-Surrogate, (2022). https://github.com/grayson-harrington/Chained-ANN-Phase-field-Surrogate.

Conflict of Interest Statement

On behalf of all authors, the corresponding author states that there is no conflict of interest.