Computational Materials Science 210 (2022) 111417

Contents lists available at ScienceDirect

Computational Materials Science

journal homepage: www.elsevier.com/locate/commatsci

Full length article

Accelerated materials design using batch Bayesian optimization: A case study for solving the inverse problem from materials microstructure to process specification

P. Honarmandi [a,*], V. Attari [a], R. Arroyave [a,b,c]

[a] Department of Materials Science and Engineering, Texas A&M University, College Station, TX, USA
[b] Department of Mechanical Engineering, Texas A&M University, College Station, TX, USA
[c] Department of Industrial and Systems Engineering, Texas A&M University, College Station, TX, USA

ARTICLE INFO

Keywords:
Microstructure descriptors
Bayesian optimization
Gaussian process
ICME framework

ABSTRACT

Microstructure-based process design is one of the main ingredients for materials design, under the integrated computational materials engineering paradigm, which relies on inverting process-structure–property linkages. The specific inverse problem connecting microstructure to processing conditions is exceedingly difficult to solve, even in a computational setting. The difficulty arises from the challenges associated with properly representing the microstructure space as well as the computational cost of the simulations used to connect process conditions to microstructure evolution. In this work, we attempt to invert a process-microstructure problem by implementing and deploying a search scheme based on multi-scale batch Bayesian optimization. We employ this framework to efficiently navigate the microstructure manifold in two examples involving phase-field simulations. In these examples, the volume fraction and characteristic length scale of phases resulting from spinodal decompositions are considered in different objective functions to find synthetic target microstructures. We show how this batch Bayesian optimization can be used to efficiently uncover process-microstructure connections through optimal parallel querying of the process space, providing a new pathway for solving inverse problems in materials design.

1. Introduction

Materials are at the forefront of new technologies, and the design/discovery of better materials enables technological innovation and economic prosperity. Goal-oriented systematic materials design, as opposed to the Edisonian approach, is central in understanding and exploiting desired properties in materials and increasing the development rate of new materials. In this regard, one of the main tasks is establishing the process-structure–property-performance (PSPP) relationship under the integrated materials computational engineering (ICME) framework. While building the forward computational relationship is straightforward due to existing modeling/simulation tools and physical knowledge about the materials systems, the backward (inverse) linkages along the PSPP chain in the context of materials design/discovery are extremely challenging. These challenges involve: (1) the model parameterization and uncertainty quantification in the presence of expensive high-fidelity models and the typical lack of data in materials science problems and (2) the absence of one-to-one mapping in the inverse problems from property/performance to microstructure to processing, meaning the same property/performance may result from multiple microstructures, as well as, the same microstructural features (descriptors) may result from multiple processing conditions in a material system.

A more noticeable lack of studies and hindering factors in inverting microstructure to processing are the main motivation of this work. Here, the extra hindering factors are mainly identified as the difficulties in the proper selection and determination of representative microstructure descriptors in these inverse (design) problems [1] and the additional complexity and cost associated with the application of the microstructure simulations. The latter factor makes the application of conventional high-throughput or trial-and-error approaches in the exploration of typically huge process space almost impractical [2–4]. This matter emphasizes the need for an efficient optimization technique to guide the spatial search toward the solution(s) with the lowest possible model evaluations.

A data-driven machine learning technique in the Bayesian framework, known as Bayesian optimization (BO) [5], has recently become popular in the materials science community to tackle the mentioned

* Correspondence to: Texas A&M University, Department of Materials Science and Engineering, 3003 TAMU, College Station, TX 77843-3003, USA.
E-mail address: hona107@tamu.edu (P. Honarmandi).

https://doi.org/10.1016/j.commatsci.2022.111417
Received 6 January 2022; Received in revised form 29 March 2022; Accepted 31 March 2022
Available online 21 April 2022
0927-0256/© 2022 Elsevier B.V. All rights reserved.

issue related to optimization efficiency and cost for the sake of accelerated materials design/discovery. Generally, this approach provides an efficient scheme to find the global optimum in the design space based on the target quantities of interest (QoIs). Unlike conventional methods, BO actively learns and updates the current state of knowledge in a sequential manner. Moreover, it searches the space with an optimal policy that keeps a balance between exploitation and exploration, rather than a non-optimal uniform search usually provided by the design of experiments (DOE) approaches, such as Latin-hypercube, sparse grid, or full grid sampling [6]. In standard BO approaches, a probabilistic surrogate model, mainly a Gaussian process (GP), is constructed over some initial input–output data obtained from the given expensive model. Then, this probabilistic information is input to an acquisition function to find the possibility of improvement as a result of the future model evaluation at each point of the input (design) space. The point that maximizes the acquisition function, i.e., the possibility of improvement, is suggested for the next model evaluation. As the surrogate model is updated by the new evaluated point, the algorithm is iterated until the model input corresponding to the target output is obtained [3,6–8].

Despite the efficiency of the standard BO approaches, the one-at-a-time selection of points restricts their practicality in the case of extremely large (high-dimensional) design space and/or super-expensive models where each model evaluation may take days or weeks. For this reason, batch BO techniques have been developed based on the computational parallelization capabilities as alternatives to accelerate the optimization process in the above cases by reaching the optimal solution with a smaller number of iterations. Different batch BO methods have been implemented during recent years that can be categorized mainly into multi-step look-ahead (hallucination) [1,9,10], pure exploration [11], and local penalization [12] techniques.

The first approach uses the standard one-at-a-time BO method in each iteration with consideration of the GP predictive mean responses at the suggested points instead of the expensive model results to obtain a batch of points from the design space for the subsequent evaluations. Although this approach is very straightforward, there is no statistical guarantee for convergence. The second category includes the variance-based and distance-based exploration approaches. The variance-based method works the same way as the multi-step look-ahead approach, except expected improvement (EI) is replaced by an acquisition function reflecting the GP predictive variance. The distance-based approach selects points with a maximum distance from their closest data point in an iterative way, while the first batch point is still suggested by the standard BO technique. Despite a slightly higher exploration efficiency of the variance-based approach, the iterative construction of the GP models makes this approach more costly than the distance-based approach, where only one GP model is built for the identification of the first point. However, these approaches neglect the spatial exploitation that can diminish their efficiency in optimization processes. In the last approach, the first batch point is again determined by the standard BO. Then, the next points are selected by maximizing the same acquisition function after excluding regions around the previously suggested points using local penalizations. This BBO approach is fast due to the construction of only one GP model. However, sometimes, the penalizer may not completely exclude the region around the maximum peak of the acquisition function, which leads to less diversity of the batch points.

Recently, Joy et al. [13] proposed a multi-scale search BBO method with none of the issues mentioned above. The main idea in this approach is to tackle the batch problem through the fundamental aspect of the GP surrogate modeling, which is the strong dependency of the surrogate accuracy on the proper selection of the length-scale hyperparameter values. Since the appropriate values are not usually determinable due to the lack of data, a large number of GP models are generated in this BBO approach by sampling the length-scale hyperparameters from their reasonable ranges. In this scenario, there will be some GP models that can approximately emulate the true response of the objective function. Therefore, at least one of the batch points resulting from clustering the points obtained by applying the standard BO method on each of these GP models might have a greater chance of being close to the global optimum of the objective function. For this reason, the multi-scale search BBO technique can be more efficient in optimization compared to the other three BBO approaches in theory. Although each iteration of this BBO approach is more expensive than its counterpart in the other approaches, the cost reduction associated with the optimization efficiency of this approach outweighs the higher cost of its optimization iterations. This becomes more evident when each model evaluation is more expensive than each BBO iteration, which seems to be the case in most materials design/discovery problems.

In the relevant literature, many works have used standard BO methods to tackle inverse problems in materials, mostly from property/performance to processing [2,6,14–28]. However, as mentioned earlier, there are very few works for microstructure-informed process design linking microstructure to processing. Tran et al.'s work [1] is the only significant study, to the best of our knowledge, in materials science and engineering that targets the inverse structure-process linkage in the PSPP chain through the application of a hallucination BBO technique. In their work, multiple batch sets are proposed through different acquisition functions to efficiently solve inverse problems across kMC simulations based on the given target microstructures represented with various descriptors [1].

In this work, the multi-scale search BBO approach is used for the efficient design of microstructures obtained through phase-field simulations. Here, the design input variables are efficiently identified given the microstructure descriptors of interest. Before solving these design problems, the better efficiency of the applied BBO method compared to the standard BO is verified using a simple sinusoidal case study. This novel framework can be the beginning of a new era in accelerated materials design/discovery due to its capability to tackle complicated inverse problems that are time-consuming or sometimes impossible to solve through conventional design approaches.

2. Model description and methodology

2.1. Phase-field modeling

The process-structure–property link is the core of materials science in which the properties of a material will be designated in the mesoscopic space due to topological and interfacial arrangements in between the existing phases. There are different methods to study the evolution of the mesoscale structure of a material. We used a coarse-grain category phase-field method in this study as a powerful tool to handle the free-boundary problems for arbitrarily complex interfaces. It describes the mesoscale state of a matter in both compositional (conserved), structural (non-conserved), and interfacial manner using a set of continuous field variables. The generality and flexibility in phase-field approach makes it convenient to use in broad range of phenomena across materials science, and beyond, including fracture [29], liquid metal dealloying [30], tumor growth studies [31], and many other applications. In this approach, the total free energy functional ($\mathscr{F}^{tot}$) for a heterogeneous solid medium as the sum of contributing fields over the domain (Ω) is:

$$\mathscr{F}^{tot} = \int_{\Omega} \left(f_{bulk} + f_{interfacial} + f_{other} \right) d\Omega \tag{1}$$

where bulk free energy, f_{bulk}, and interfacial free energy, $f_{interfacial}$ are given below, respectively, and f_{other} represents nonlocal contributions from any long-range interactions, (e.g., elastic, electric, magnetic effects, etc.). The main differences among different phase-field models lie in the treatment of various contributions to the total free energy.

$$f_{bulk} = f^0\left(c(\boldsymbol{r}), \eta(\boldsymbol{r}), T\right), \tag{2}$$

$$f_{interfacial} = \frac{1}{2}\kappa_c |\nabla c(\boldsymbol{r})|^2 + \frac{1}{2}\sum_p \kappa_{ij}^{\eta} \nabla_i \eta_p(\boldsymbol{r}) \nabla_j \eta_p(\boldsymbol{r}), \tag{3}$$

where c is the composition field (conserved), η is the non-conserved field order parameter, $\boldsymbol{r}$ is the position vector, $f^0(c,\eta,T)$ is the free energy of a unit volume of homogeneous material for a given temperature (T), κ_c is the composition-gradient energy coefficient, and κ_{ij}^{η} is the symmetric tensor of field-order-gradient energy coefficients. With the total free energy of a microstructure discussed above, the evolution of field variables in a phase-field model can be obtained by solving the Cahn–Hilliard (4) and Allen–Cahn (5) equations,

$$\frac{\partial c_i(\boldsymbol{r},t)}{\partial t} = \nabla M_{ij} \nabla \frac{\delta \mathscr{F}^{tot}}{\delta c_j(\boldsymbol{r},t)} \tag{4}$$

$$\frac{\partial \eta_p(\boldsymbol{r},t)}{\partial t} = L_{pq} \frac{\delta \mathscr{F}^{tot}}{\delta \eta_q(\boldsymbol{r},t)} \tag{5}$$

where M_{ij} and L_{pq} are related to atom or interface mobility, and δ is the variational derivative operator.

Two phase-field models are considered in this work for developing the inverse PSP framework. Here, we briefly describe these models. A (i) low-fidelity, benchmark-type 2D model for a simple case of phase separation that only involves chemical and interfacial interactions is initially used to develop the framework, and later (ii) a high-fidelity 3D elasto-chemical phase-field model that also considers elastic interactions during phase separation in a ternary alloy is used. These models are described as follows:

Low-fidelity phase-field model. A chemical-only phase-field model based on the model of Cahn–Hilliard (C–H) [32] is initially used for developing the framework as well as testing our BBO approach. We consider this basic C–H model as an inexpensive low-fidelity model since it only considers chemical and curvature effects. The model treats the process of spontaneous phase separation in an incompressible isotropic solution of constant molar volume. The boundary value problem is given as:

$$\begin{aligned} &\frac{\partial \phi}{\partial t} = -M\nabla^2\left(\mu - \kappa_\phi \nabla^2 \phi\right) \quad 0<x<L_x \quad 0<y<L_y \quad 0<t<t^* \\ &BC: \phi(0,y,t) = \phi(L_x,y,t) \quad \phi(x,0,t) = \phi(x,L_y,t) \\ &IC: \phi(x,y,0) = \phi^* + A\zeta \end{aligned} \tag{6}$$

where ϕ is a field order parameter, $\kappa_\phi = 0.005$ is the gradient energy coefficient, and $M = 1$ is the inherently positive effective rate of atomic mobility for the species. For simplicity, the values of M and κ_ϕ are fixed. The lengths of the simulation domain are set to $L_x = L_y = 2\pi$ with grid size of 256×256 and t^* is the final model run time. BC and IC denote the used boundary and initial conditions, respectively. ϕ^* is the initial average value of the order parameter that is perturbed by a constant noise magnitude A, and ζ is a Gaussian random number with the interval of $[-1, +1]$.

By using the Fourier-spectral approximation to C–H equation, and, treating the linear fourth-order operators implicitly and the nonlinear terms explicitly, the first-order semi-implicit Fourier-spectral scheme [33,34] is used to solve this equation. The resulting first-order semi-implicit Fourier spectral scheme is:

$$(1 + M\Delta t \kappa k^4)\tilde{\phi}^{n+1}(\mathbf{k}) = \tilde{\phi}^n(\mathbf{k}) - \Delta t, k^2\{\tilde{\mu}(\phi^n)\}_{\mathbf{k}} \tag{7}$$

where $\mathbf{k} = (k_1, k_2)$ is a frequency vector in the Fourier space, $k = \sqrt{k_1^2 + k_2^2}$ is the magnitude of this vector and $\tilde{\phi}(\mathbf{k},t)$ represents the Fourier transform of $\phi(\mathbf{r},t)$. μ is the driving force for phase transformation and is given as the first derivative of the nonlinear term (bulk free energy):

$$\mu = \frac{\partial f(\phi)}{\partial \phi} \quad \text{where} \quad f(\phi) = -\frac{1}{2}\phi^2 + \frac{1}{4}\phi^4. \tag{8}$$

This free-energy function has two degenerate minima represented at $\phi = -1$ and $\phi = +1$ corresponding to the products of the process of phase separation. ϕ^* and t^* will be determined by the BBO method during the optimization process, given a synthetic target microstructure shown in Fig. 1(a).

High-fidelity phase-field model. In our second case study, we illustrate the microstructure-based inverse problem using an elasto-chemical phase-field that attributes the prescribed elastic contractions to solid-state phase transitions in addition to the chemical and interfacial interactions. In addition, the bulk free energy term is a CALPHAD form free energy function optimized by experimental equilibrium data for the ternary alloy under study. Accordingly, the thermodynamic driving force for microstructural evolution is associated with chemical interactions resulting from compositional fluctuations and local contractions/expansions caused by composition- and temperature-dependent elastic constants and lattice parameters. Furthermore, in this section we perform 3D simulations to test the capability of the framework with real structures. We study the TiAlZrN alloy that is known to decompose into its stable binary nitrite compounds (i.e., TiN, AlN, ZrN) during isothermal thermal annealing and is thermodynamically treated as a pseudoternary system. This system has been previously investigated by some of the authors in this article and the results were published in [35]. The evolution of the conserved order parameters (c_α) are given by:

$$\begin{aligned} &\frac{\partial c_\alpha}{\partial t} = \nabla.\left[M_\alpha \nabla\left(\frac{\partial f_{bulk}(c_\alpha)}{\partial c_\alpha} + \mu_{elastic} - \kappa_\alpha \nabla^2 c_\alpha - \Sigma_{\beta=1}^{N} L_{\alpha\beta}\nabla^2 c_\beta\right)\right] \\ &0<x,y<L \quad 0<z<L_z \quad 0<t<t^* \\ &BC: c(0,y,z,t) = c_\alpha(L,y,z,t) \quad c_\alpha(x,0,z,t) = c_\alpha(x,L,z,t) \\ &\quad c_\alpha(x,y,0,t) = c_\alpha(x,y,L_z,t) \\ &IC: c_\alpha(x,y,z,0) = c_\alpha^0 + A\zeta \end{aligned} \tag{9}$$

For the ternary alloy, the system of kinetic equations for the evolution of the composition of species (α = AlN, ZrN, TiN) follows three sets of kinetic equations. However, since the composition is a conserved order parameter and the relationship of ($c_{\mathrm{TiN}} = 1 - c_{\mathrm{AlN}} - c_{\mathrm{ZrN}}$) holds, we only solve two equations for AlN and ZrN. The bulk free energy f_{bulk}, has the form of a CALPHAD free energy given below:

$$f_{bulk} = \sum_{\alpha}^{N} c_\alpha G_\alpha^o + RT\sum_{\alpha}^{N} c_\alpha \ln(c_\alpha) + \sum_{\alpha}^{N}\sum_{\beta>\alpha}^{N} c_\alpha c_\beta \sum_{\nu=0}^{\lambda} L_{\alpha\beta}^{\nu}(c_\alpha - c_\beta)^\nu + G^{xs} \tag{10}$$

where G_α^o is the reference Gibbs energy, R is the ideal gas constant, and $L_{\alpha\beta}^{\nu}$ are the excess binary interaction parameters, and $\lambda = 1$ represents a sub-regular thermodynamic system. G^{xs} is a higher order interaction parameter defined as:

$$G^{xs} = \sum_{\alpha}^{N-2}\sum_{\beta>\alpha}^{N-1}\sum_{\gamma>\beta}^{N}\left(c_\alpha c_\beta c_\gamma L_{\alpha\beta\gamma}\right) \tag{11}$$

where $L_{\alpha\beta_\gamma}$ is the excess ternary interaction parameter taken to be $L_{\alpha\beta\gamma} = {c_\alpha}^{\alpha} L_{\alpha\beta\gamma} + {c_\beta}^{\beta} L_{\alpha\beta\gamma} + {c_\gamma}^{\gamma} L_{\alpha\beta\gamma}$. In this system, none of these interaction parameters are zero. All these thermodynamic parameters are obtained from [36,37] and are summarized in Table 1 (the thermodynamic parameters used in the free energy model) of [35].

The elastic potential ($\mu_{elastic} = \frac{df_{elas}}{dc}$) will be obtained by expressing the elastic free energy in linear elasticity theory as:

$$f_{elas} = \frac{1}{2} C_{ijkl}\left(E_{ij} + \varepsilon_{ij}^* - \varepsilon_{ij}^0\right)\left(E_{kl} + \varepsilon_{kl}^* - \varepsilon_{kl}^0\right) \tag{12}$$

where C_{ijkl} is the elastic stiffness tensor, E_{ij} is the average homogeneous strain tensor in the domain Ω, ε_{ij}^* is the periodic fluctuation strain field ($\varepsilon_{ij}^{tot} = E_{ij} + \varepsilon_{ij}^*$), and ε_{ij}^0 is the dilatational eigenstrain derived from the lattice misfit. Since the composition field is periodic, the elastic moduli and the eigenstrains are thus periodic on the material domain Ω. Therefore, the total strain field ε_{ij}^{tot} is periodic and can be expressed as the sum of E_{ij} and ε_{ij}^*. The elastic stiffness tensor is obtained by interpolating across the simulation unit cell through the interpolation function $g(c_\alpha)$ taken from [38]:

$$C_{ijkl}(c) = C_{ijkl}^{eff} + \sum_{\alpha=1}^{N-1} g(c_\alpha)\left(C_{ijkl}^{\alpha}(T) - C_{ijkl}^{0}(T)\right) \tag{13}$$

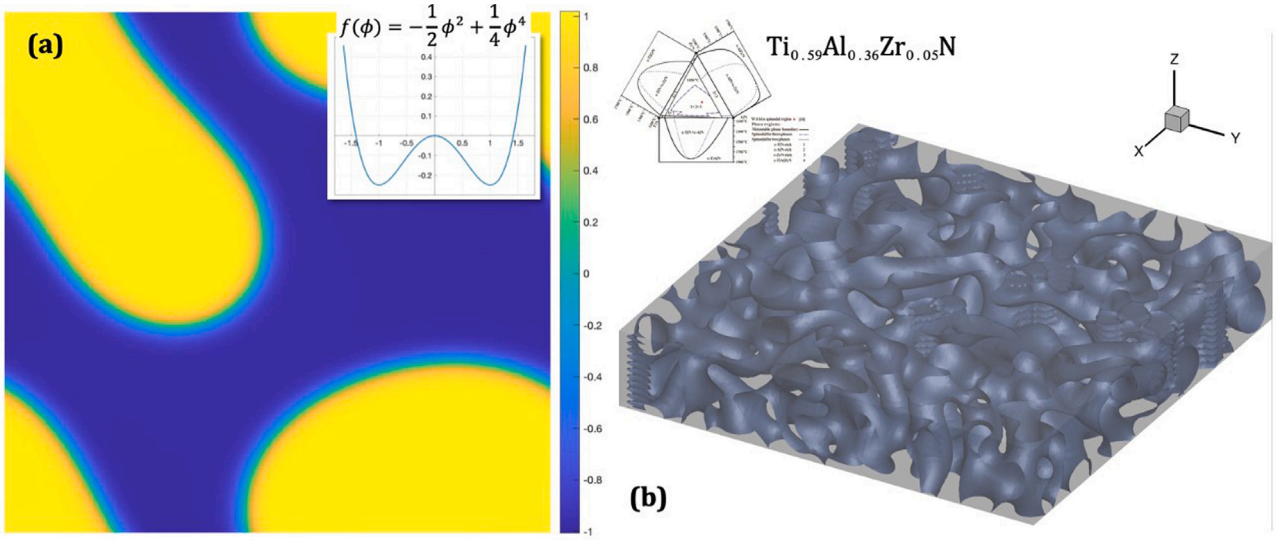

Fig. 1. (a) 2D representative microstructure obtained from the low-fidelity model for the target input values of $(\phi^*, t^*) = (0, 30)$, and (b) 3D nanostructure of TiAlZrN alloy obtained from the high-fidelity model for the target input values of $(c^*_{\mathrm{AlN}}, c^*_{\mathrm{ZrN}}, t^*) = (0.36, 0.05, 38)$. The isosurface contours show the bicontinuous TiN-AlN interfaces where small amounts of ZrN phase segregate.

where C^{eff}_{ijkl} represents the effective elastic stiffness tensor and is selected to be the average of the elastic tensors of the product phases of the spinodal decomposition. C^0_{ijkl} and C^α_{ijkl} represent the elastic tensor of the reference phase and the other phases, respectively. The mechanical equilibrium equation given below [35]:

$$\begin{cases} \frac{\partial}{\partial r_j}\left\{C_{ijkl}\left(E_{kl} + \varepsilon^\star_{kl} - \varepsilon^0_{kl}\right)\right\} = 0 & \text{on} \quad \Omega \\ \varepsilon^\star_{kl} & \text{periodic on } \Omega \end{cases} \tag{14}$$

is efficiently solved by transforming it into the *Fourier* space and solving it through the standard FFT based iterative solvers described by Khachaturyan [39], Mura [40], Moulinec-Suquet [41], Gururajan et al. [42], Lebensohn et al. [43], and our previously published article [35]. On the other hand, the numerical solution of the boundary value problem shown in Eq. (9) was obtained by the semi-implicit *Fourier* spectral scheme developed in [35]. For each iteration (i.e., at each time step of the phase-field solver), we solve the microelasticity problem for a given composition at a given time. The size of the 3D simulation cell is chosen to be $50 \times 50 \times 11.9$ nm^3, and it is discretized with 64 × 64 × 16 mesh points. The simulations were performed at 1200 °C and all temperature-dependent elastic parameters for the TiAlZrN system were obtained from [35] and enclosed references in this article. Except c^*_α and t^*, all other parameters are fixed in this model and c^*_α and t^* will be determined by the BBO framework during the optimization process, given a synthetic target microstructure shown in Fig. 1(b).

2.2. Microstructure representation

One challenge in the inverse design of material microstructure is the compact representation of microstructure images for effective optimization of material and/or process variables. An image of the microstructure containing bulk and interfacial features often reflected in different color thresholds is a valuable source of information about the material. Consequently, every pixel of the image represents a degree of freedom, causing the dimensionality of the information space to be extremely high. There have been numerous studies on how to reduce high-dimensionality and represent images properly. A brief review is provided here [44]. In addition to the conventional physical descriptors (e.g., area/volume fraction of phases), other features have been proposed to aggregate local information into global descriptors. Some of these include, but are not limited to, Haralick texture statistics [45], and other contrast features, as well as local binary patterns. There is also a method known as Visual Bag of Words [46], which constructs a dictionary of commonly occurring visual keypoints and stores a descriptor for an image by counting the occurrence of each of these commonly occurring visual keypoints. A different but widely used method is the auto- and cross-correlation function (e.g., two-point correlations) of a binary image. An example application of two-point correlations can be found in [47] and also articles published by a portion of the authors in this article [48,49].

In this study, we avoid further complicating the problem and aim to build the framework using conventional physical descriptors (i.e., area fraction for (2D) calculations and volume fraction for (3D) calculations, both indicated by V_f, and Fourier scattering intensity of microstructures) that are readily accessible to metallurgists. Indeed, more complex microstructure image representation methods can be embedded when necessary. V_f of phases in 2D simulations is estimated by computing a global image threshold using Otsu's method [50] from grayscale images. The method works well for semi- and fully-decomposed microstructures where the area fractions of the phases are approximately greater than 0.2. However, it cannot be used at the early stages of decomposition when the structure is still uniform. The volume fraction of phases in 3D calculations is estimated by phase composition thresholding over grid points. Fig. 2(a) shows the calculated area fraction of phases as a function of time for the case where the average field order parameter at time zero (ϕ^*) is zero. The final point at $t^* = 30$ is the designated 2D target microstructure shown in Fig. 1(a). In Fig. 2(b), the change in 2D microstructures corresponding to the annealing duration of $t^* = 30$ as a function of ϕ^* is depicted. As indicated, V_f of yellow phases increases as the value of ϕ increases from −0.4 to +0.4 when annealing time is fixed in a binary system. Fig. 2 indicates that the area/volume fraction is not unique for each microstructure in time and ϕ^* space and additional quantities of interest are necessary to identify each unique microstructure.

The second quantity of interest that we use is calculated by means of circularly averaged Fast Fourier Transform (FFT) of the field order parameters (i.e., low fidelity model: $\phi(x, y)$, high-fidelity model: $c^\alpha(x, y, z)$). Fig. 3(a) shows the circularly averaged magnitude of the FFT as a function of time for a 2D simulation. FFT is shifted to the center for both 2D and 3D field order or composition arrays. The calculations show a unique curve for each image at a fixed time where the curve and

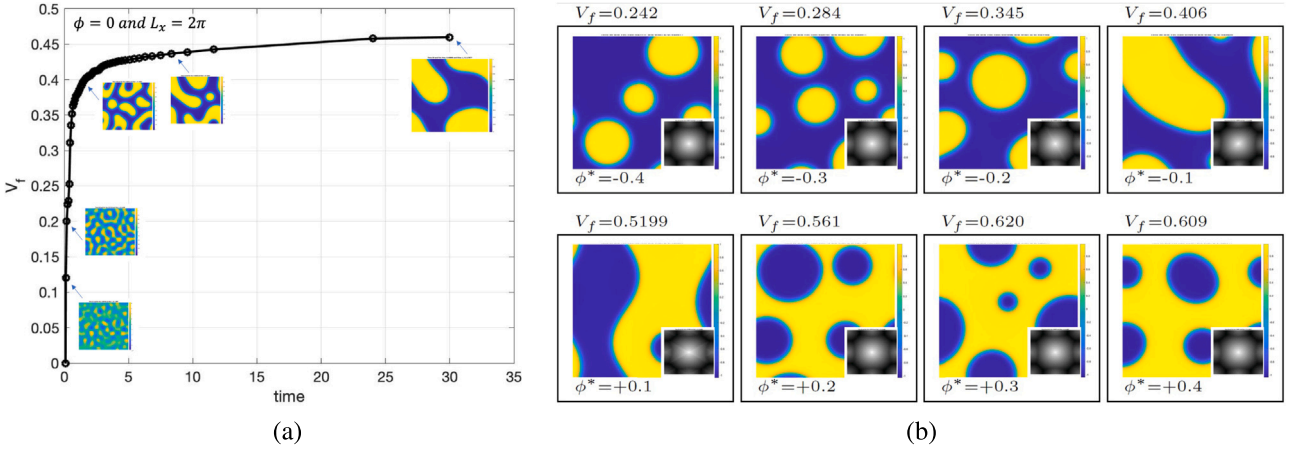

Fig. 2. (a) The evolution of area fraction (V_f) of yellow phases as a function of time for system with $\phi^* = 0$. Several microstructures corresponding to different times are shown next to the curve; (b) The corresponding Change in morphology as a function of a change in ϕ^* after a fixed annealing duration ($t^* = 30$). Area fraction of yellow phases increases when ϕ^* increases from −0.4 to +0.4, i.e., moving from left hand side of the spinodal region to the right hand side over an isothermal line. Shifted power spectral density (PSD) or power spectrum of microstructures are shown in the right corner of each image.

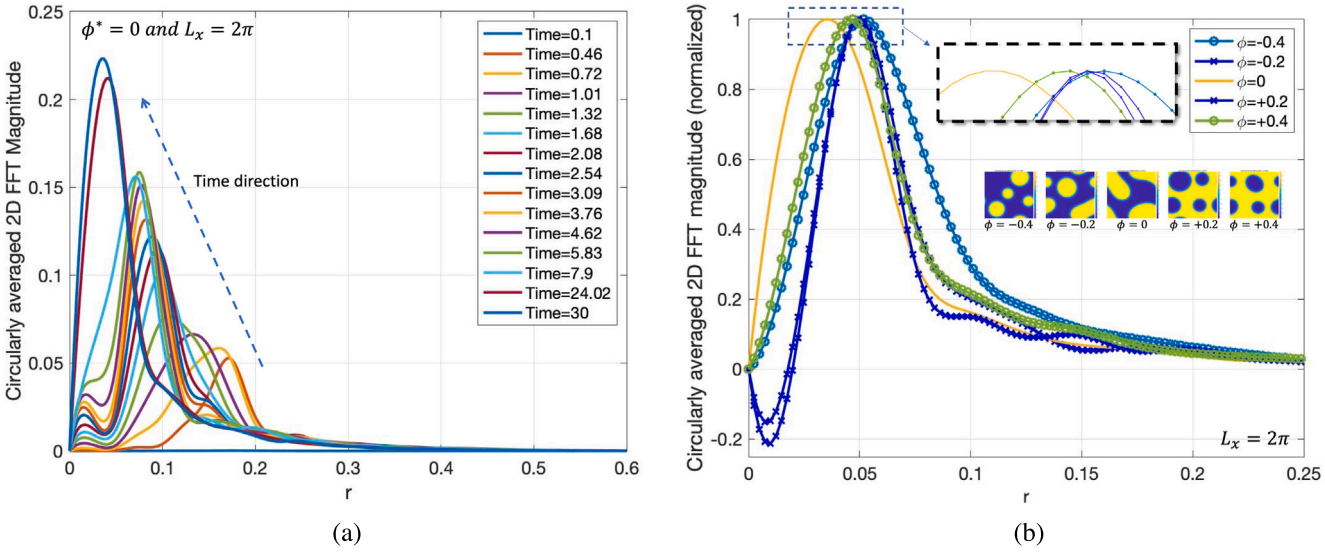

Fig. 3. (a) Circularly averaged magnitude of FFT for the simulation with $\phi^* = 0$ and $L_x = 2\pi$, as a function of annealing time. The curves and their peak shifts to the right as the annealing time increases; (b) Circularly averaged magnitude of FFT as a function of a change in value of the average phase-field order parameter (ϕ^*) showing small differences in the peak locations enabling us to distinguish the small microstructures.

its peak shifts to the right-hand side as the annealing time increases. The corresponding frequency to the peak in these curves is extracted and used to calculate a length scale (L_s). Likewise, the circularly averaged magnitude of the FFT for different microstructures corresponding to different ϕ^* values annealed equally until $t^* = 30$ is shown in Fig. 3(b). While the L_s value is close in these cases, the zoomed section of the peak area shows that there are small differences in the peak location corresponding to each microstructure. Later on in this paper, we show that the consideration of length scale besides area/volume fraction can help for the unique representation of microstructures.

2.3. Design methodology: Multi-scale search batch Bayesian optimization

Analogous to the standard BO, the Multi-scale Search BBO method takes advantage of the data-driven machine learning (ML) models, specifically GP surrogate models, as substitutes for expensive models to accelerate the optimization process in computational materials design. GPs are non-parametric stochastic ML models in the context of the Bayesian statistics that consider a multivariate Gaussian distribution for any given sample ($\{f(\mathbf{x}_1), f(\mathbf{x}_2), \ldots, f(\mathbf{x}_n)\}$) from the objective function $f(\cdot)$. This Gaussian distribution over the functions is defined as $f(\mathbf{x}) \sim \mathrm{GP}(\mu(\cdot), k(\cdot,\cdot))$, where $\mu(\cdot)$ and $k(\cdot,\cdot)$ are the mean function and covariance (kernel) function, respectively. The kernel function is commonly expressed in the squared exponential (radial basis) form for any given pair points in the space, such as $\mathbf{x}$ and $\mathbf{x}'$, as follows:

$$k(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{1}{2}\sum_{i=1}^{d}\frac{(x_i - x_i')^2}{l_i^2}\right) \tag{15}$$

where d is the dimension of the input space, and $\mathbf{l}$, known as the length-scale hyperparameter, is a $1 \times d$ vector that determines the smoothness of the predictive function.

Given the data corresponding to the objective function, i.e., $\mathbf{D} = \{\mathbf{x}_r, y_r\}_{r=1}^{R}$ such that $y_r = f(\mathbf{x}_r) + \epsilon$ and $\epsilon \sim \mathcal{N}(0, \sigma^2)$, the GP prior with a typically zero mean function is updated to a GP posterior by maximizing the marginal likelihood with respect to the GP hyper-parameters, i.e., $\theta = \{\sigma, \mathbf{l}\}$. Therefore, a posterior predictive distribution can be obtained at any given spatial point $(\mathbf{x}_{R+1}, \mathbf{y}_{R+1})$ based on the existing joint Gaussian distribution between $\mathbf{y}_{1:R}$ and y_{R+1}, as follows,

$$\begin{bmatrix}\mathbf{y}_{1:R} \\ y_{R+1}\end{bmatrix} \sim \mathcal{N}\left(0, \begin{bmatrix}\mathbf{K} + \sigma^2\mathbf{I} & \mathbf{k} \\ \mathbf{k}^T & k(\mathbf{x}_{R+1}, \mathbf{x}_{R+1}) + \sigma^2\end{bmatrix}\right) \tag{16}$$

$$p(y_{R+1}|\mathbf{x}_{R+1},\mathbf{D}) \sim \mathcal{N}\left(\mu_{pred}(\mathbf{x}_{R+1}),\sigma^2_{pred}(\mathbf{x}_{R+1})\right) = \mathcal{N}\left(\mathbf{k}^T[\mathbf{K}+\sigma^2\mathbf{I}]^{-1}\mathbf{y}_{1:R}, k(\mathbf{x}_{R+1},\mathbf{x}_{R+1})+\sigma^2 - \mathbf{k}^T[\mathbf{K}+\sigma^2\mathbf{I}]^{-1}\mathbf{k}\right) \quad (17)$$

where $\mathbf{K} = \begin{bmatrix} k(\mathbf{x}_1,\mathbf{x}_1) & \dots & k(\mathbf{x}_1,\mathbf{x}_R) \\ \vdots & \ddots & \vdots \\ k(\mathbf{x}_R,\mathbf{x}_1) & \dots & k(\mathbf{x}_R,\mathbf{x}_R) \end{bmatrix}$, $\mathbf{k} = \{k(\mathbf{x}_1,\mathbf{x}_{R+1}), k(\mathbf{x}_2,\mathbf{x}_{R+1}), \dots, k(\mathbf{x}_R,\mathbf{x}_{R+1})\}$, and $\mathbf{I}$ denote the $R\times R$ identity matrix [13,51].

As mentioned earlier, the length-scale hyperparameter that controls the smoothness of the predictive surrogate surface is an important factor to emulate the response of the black-box objective function. However, the correct estimation of this parameter is usually impossible due to the lack of data for the objective function. For this reason, Joy et al. [13] incorporated the idea of building lots of surrogate models by sampling the length-scale hyperparameter from a reasonable range of values into the standard BO technique to propose a multi-scale search BBO approach. Considering this strategy in the BO scheme ensures that some of the constructed surrogate surfaces are close to the response of the objective function, which makes this BBO approach highly robust and super-efficient.

Assuming G length-scale samples ($\{\mathbf{l}_g\}_{g=1}^G$) and the corresponding G surrogate models ($\text{GP}_{g=1}^G$), then G points ($\{\mathbf{x}_g^*\}_{g=1}^G$) can be suggested using the maximization of the acquisition function (α), as follows,

$$\mathbf{x}_g^* \triangleq \arg\max_{\mathbf{x}\in\chi} \alpha\left(\mathbf{x}|\text{GP}_g(\mathbf{D},\mathbf{l}_g)\right) \quad \forall g\in\{1,\dots,G\} \quad (18)$$

Expected improvement (EI) is considered as the acquisition function in this work in order to hold a balance between exploitation and exploration of the optimization space. It is expressed as the expected value of the improvement (I) at any given spatial point ($\mathbf{x}$), as follows [5],

$$\alpha_{EI}(\mathbf{x}) \equiv E[I(\mathbf{x})|\mathbf{x},\mathbf{D}] = E[\max(y_{min}-y(\mathbf{x}),0)] = (y_{min}-\mu_{pred}(\mathbf{x}))\Phi\left(\frac{y_{min}-\mu_{pred}(\mathbf{x})}{\sigma_{pred}(\mathbf{x})}\right) + \sigma_{pred}(\mathbf{x})\phi\left(\frac{y_{min}-\mu_{pred}(\mathbf{x})}{\sigma_{pred}(\mathbf{x})}\right) \quad (19)$$

which is an expression specifically for the minimization cases. Here, $y_{min} = \arg\min_{y_r\in\mathbf{y}_{1:\mathbf{R}}} y_r$ is the minimum y value of the current data points. The second line in Eq. (19) shows a closed analytic solution for the expected value in the case of single-objective problems where the objective function response is one-dimensional. $\Phi(\cdot)$ and $\phi(\cdot)$ denote the standard normal cumulative distribution function (CDF) and probability distribution function (PDF), respectively. In this closed-form expression, the first term provides the exploitation search while the second term accounts for the exploration search.

The G points proposed through the above process are reduced to the batch size of interest by clustering the points into K groups through the K-medoids clustering method (see [52] for the method description). In this context, the medoid of each cluster—that is an actual point among the proposed points (unlike the centroid in the K-means clustering method, which corresponds to a point in the design space that may not correspond to an actual sample)—is taken into account as the next point to evaluate.

$$\mathbf{x}_k^m = \text{K-medoids}\left(\{\mathbf{x}_g^*\}_{g=1}^G, K\right) \quad (20)$$

$\mathbf{D}$ is updated by the new data points obtained from the parallel evaluations of the above-suggested points, i.e., $\mathbf{D} = \left[\{\mathbf{x}_r,y_r\}_{r=1}^R \cup \{\mathbf{x}_k^m,y_k^m\}_{k=1}^K\right]$; then, the next optimization iteration is performed in the same way with the updated data. The optimization process continues in this sequential manner until a point with the optimum response is identified [13]. The flowchart in Fig. 4 also shows how this BBO approach works.

3. Results and discussion

3.1. Standard vs. batch Bayesian optimization: A 2D sinusoidal case study

Before applying the multi-scale search BBO technique in a practical design problem, the efficiency of this approach is compared with the standard BO through a simple sinusoidal example to show the importance of the BBO in the accelerated materials design/discovery when the parallel computing capability is available.

Let $y(\mathbf{x}) = \sum_{i=1}^2 x_i\sin(x_i)+50$ be the objective function with 2D input ($\mathbf{x}=(x_1,x_2)$) and 1D output (y). Fig. 5(a) shows a 3D demonstration of how this function changes over the input space ranging from 0 to 20 in each dimension. As can be observed in this figure, there are several local minima that make finding the global minimum difficult through conventional techniques such as gradient-based optimization methods. The color 2D projection in Fig. 5(b) shows this global minimum in dark blue around the coordinate (17.4, 17.4).

The standard BO and the multi-scale search BBO with different batch sizes are applied to find the global optimum of the above sinusoidal function, i.e., $\arg\min_{\mathbf{x}} y(\mathbf{x})$. For these optimization processes, the posterior predictive distributions and the acquisition function values are obtained at a number of discrete coordinates over the $\mathbf{x}$-input space, $\{(x_1,x_2)|x_{i=1,2}\in\{0,0.5,1.0,\dots,19.5,20\}\}$. In this discrete space, the closest point to the global optimum is located at (17.5,17.5) that can be assumed to be the optimization solution. However, the range of $(17,17)\le\mathbf{x}\le(18,18)$, which includes 9 points in the discrete input space, is considered as the optimization stopping criterion for the sake of time reduction. It should be noted that 100 random three-point sample sets from the input space excluding the mentioned objective range are selected as the initial data and used one by one for each optimization approach.

The mean value and standard deviation of the number of iterations required to achieve the optimal result for each optimization approach are plotted in Fig. 6. As observed in this figure, the standard BO with one-at-a-time evaluations has a higher mean number of iterations to the optimal result compared to the BBO with more than one-at-a-time evaluations. Regardless of the stochastic nature of the applied BBO approach causing small fluctuations in the mean number of iterations, there is an overall decreasing trend for this mean value as the optimization batch size increases. However, this reduction has a sharp decreasing rate, such that there are almost no differences in the mean number of iterations at batch sizes larger than 20. Interestingly, while the mean number of iterations for the BBO with batch size 2 is slightly smaller than the standard BO, its corresponding standard deviation is much larger. Similar to the mean number of iterations, the standard deviation also decreases with a sharp decreasing rate as the BBO batch size increases. Again, the stochastic nature of the applied BBO approach results in some small fluctuations in the standard deviation trend, but they are ineffective on the overall decreasing trend.

3.2. Application of the multi-scale search batch Bayesian optimization in inverse structure-process problems

In this section, the multi-scale search BBO method is applied to two different phase-field models explained in detail in Section 2.1 in order to solve the inverse structure-process problems in the context of accelerated materials design. In this regard, the BBO technique described in Section 2.3 is used here to search composition-time input spaces for the given target microstructures. Besides the application of the BBO technique, the importance of the proper definition of the objective function in terms of microstructure descriptors and the proper weighting of these descriptors is also demonstrated in these phase-field studies.

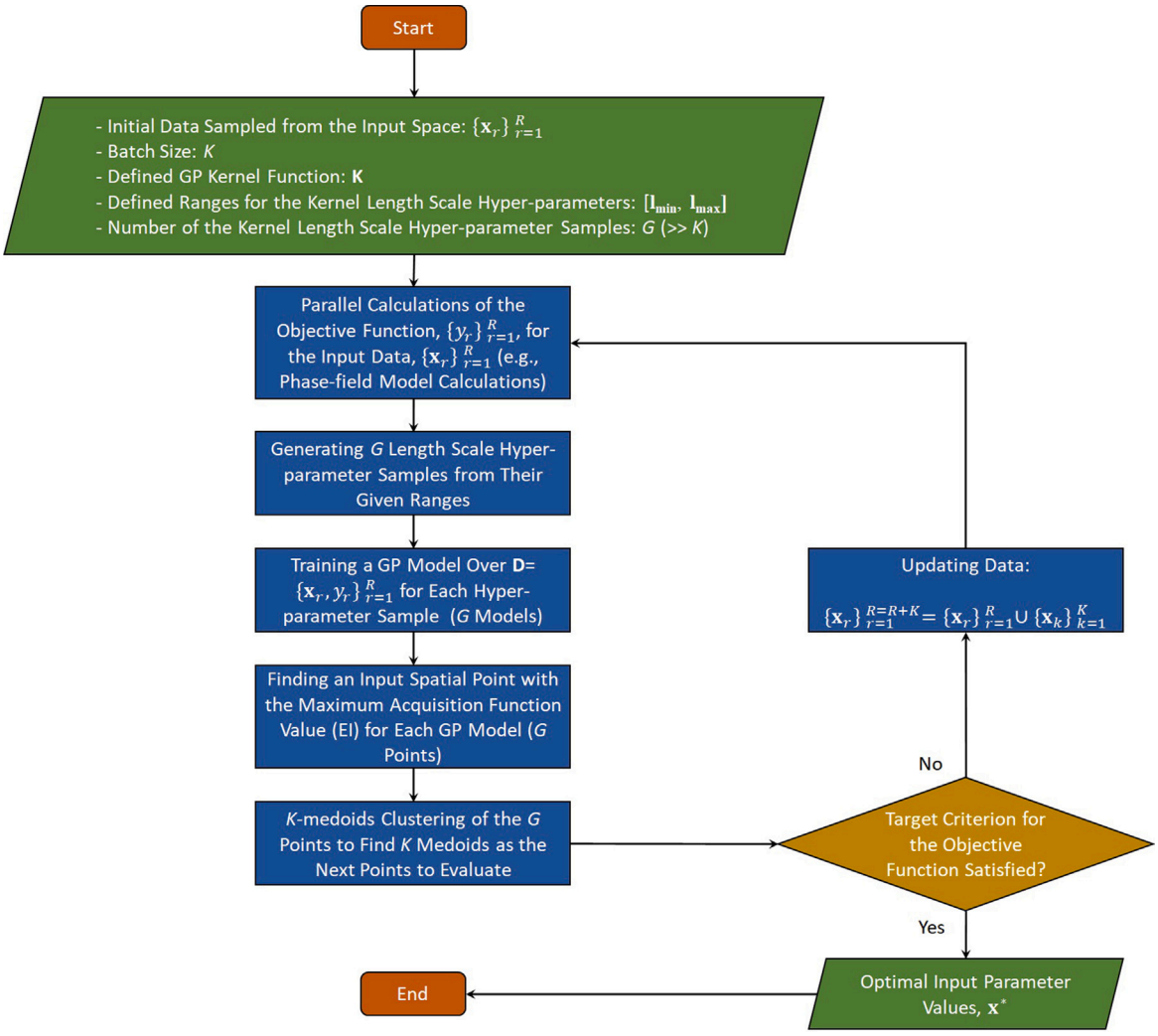

Fig. 4. A flowchart representing the multi-scale search batch Bayesian optimization.

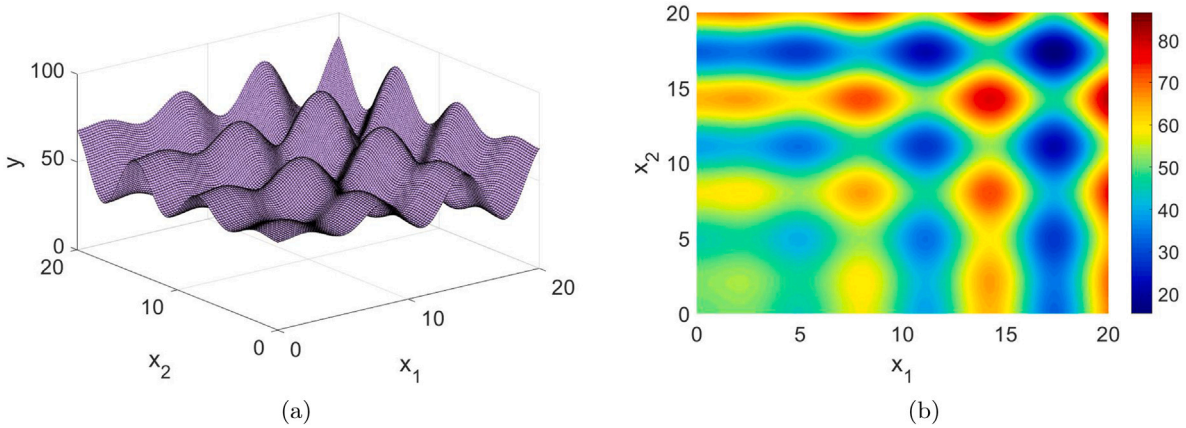

Fig. 5. (a) The response of $y(\mathbf{x}) = \sum_{i=1}^{2} x_i \sin(x_i) + 50$ over the input space ranging from 0 to 20 in each dimension ($0 \leq x_{i=1,2} \leq 20$); (b) The color 2D projection of $y(\mathbf{x})$ on the $\mathbf{x}$-input space. The global minimum around the coordinate (17.4,17.4) is illustrated in dark blue.

3.2.1. Computational microstructure-based process design using the low-fidelity model

Process design is performed by considering a synthetic microstructure generated through the low-fidelity phase-field simulation as the design target. The BBO method with different batch sizes are applied to identify the model input values, i.e., order parameter (ϕ) and time (t), which deliver the target microstructure. Here, the target input values are $(\phi^*, t^*) = (0, 30)$ and the respective microstructure is shown in Fig. 1(a). Two objective functions are defined for these optimization processes. The first objective function includes just one of the microstructure descriptors, where the volume fraction of the spinodal phases at any given point in the design space are compared with their

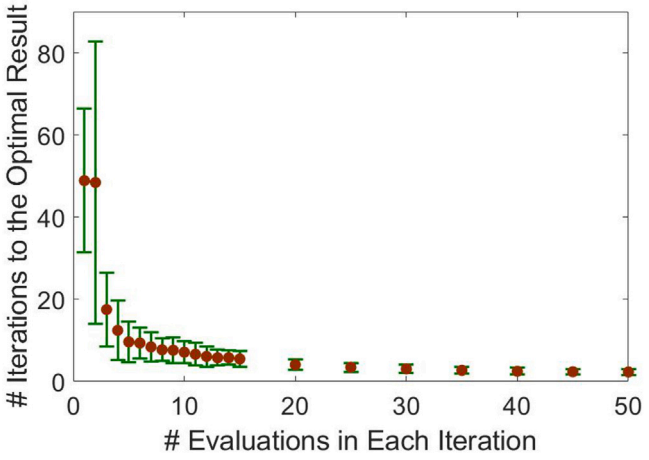

Fig. 6. The probabilistic results for the number of iterations required to achieve a point in the optimal range versus the number of evaluations in each optimization iteration. It includes the results for the standard BO with one-at-a-time evaluations and the multi-scale search BBO with different numbers of evaluations at-a-time based on the applied batch sizes.

counterparts at the target point in the form of $f_1 = |V_f - V_f^*|$. It should be noted that the volume fraction of just one phase is taken into account in the optimizations due to the correlation between the volume fraction of the two phases. However, the second objective function contains the characteristic length scale of phases besides their volume fractions and compares these two microstructure descriptors at any given point in the design space with their counterparts at the target point as $f_2 = |V_f - V_f^*| + |L_s - L_s^*|$. Again, the correlation between the descriptors of the spinodal phases justifies an objective function with the descriptors of just one of the phases. The phase volume fractions are calculated by incorporating Otsu's image thresholding method [50]. Their characteristic length scale is obtained by calculating the radial basis power spectrum of the Fourier transform of the composition map ($\log|\{c_k\}|^2$) for 2-D and 3-D microstructures. These two sets of minimization problems with different objective functions are initialized with three objective-function values obtained by the phase-field simulations at three same data points randomly sampled from a discrete input space, where ϕ and t change from −0.5 to +0.5 (based on the predefined composition range for the spinodal decomposition in the total Gibbs free energy) and 10 to 300 with the interval of 0.02 and 5, respectively. These data points are used to train the GP models used in the first iteration of the optimizations. The rest of the spatial points are considered as the test points at which the posterior predictive distributions and the acquisition function values are obtained to propose the next points to evaluate. It is also worth noting that 1000 length-scale hyperparameter samples of size 2 are generated in order to perform these batch optimizations, according to Section 2.3.

In Table 1, the numbers of the optimization iterations required to reach the target input (n^*) are listed for different optimization batch sizes and objective functions. As expected, there is a decreasing trend for the number of iterations as the batch size increases, regardless of the choice of the objective function. However, the decreasing trend can be identified more accurately if the average number of iterations after the execution of several BBOs with different initial samples is considered for each batch size case, similar to the evaluations in Section 3.1. These analyses are avoided here since the phase-field model involvement in the mentioned optimizations for each batch size case makes the whole process computationally very expensive. Moreover, the main points in this section are to show how to efficiently perform microstructure-based process design from a random initial data set in the design space and how the computational parallelization capability can considerably accelerate design by reducing the number of iterations during the optimization process.

Table 1
The number of optimization iterations required to obtain the target input using the low-fidelity phase-field model, considering different optimization batch sizes and two objective functions.

Batch size (K)	Number of optimization iterations (n^*)	
	Objective-1 (f_1)	Objective-2 (f_2)
10	78	17
20	37	9
30	19	7
40	12	5
50	9	3
100	5	2

Another important point observed in Table 1 is that the number of iterations to the design solution for each batch size case is much lower when the second objective is taken into account. This can be attributed to the fact that the volume fraction of phases alone is not generally a sufficient feature for the unique representation of microstructures. However, the volume fraction and characteristic length scale of phases together provide a better representation for microstructures in this design case with the low-fidelity phase-field Model.

Fig. 7 provides a better understanding of why objective function f_2 is more informative than objective function f_1 in this case study. In this figure, the values of both objective functions at points suggested by the BBOs with batch size 10 have been plotted in terms of field order parameter and time—where the colors show the points suggested in different optimization iterations for each objective function case. It should be noted that the three initial points for optimizations are also included in this figure in a different color. As observed in Fig. 7(a) and (b), there are many combinations of order parameter and time that results in objective function values very close to zero. Therefore, the BBO approach most likely spends some time searching the local minima until it finds the global optimum. On the other hand, as shown in Fig. 7(c) and (d), the global optimum is more defined and differentiated from the rest of the spatial points for the second objective case. For this reason, the global optimum can be identified much faster by the BBO approach. In this regard, Hardy and Voorhees [53] showed that the volume fraction and characteristic length scale are correlated at any given time, indicating if one is known, the other one can be calculated. It was also indicated that the same value of volume fraction can be obtained at different times, differentiated by different length scale values. This is the case in the present work, where just the volume fraction of phases in objective function f_1 results in solution degeneracy, i.e., objective values very close to zero for different microstructures obtained from different combinations of order parameter and time (Fig. 7(a) and (b)). This solution degeneracy can also be inferred through Fig. 2. It should be noted that the same story exists for the characteristic length scale of phases alone, based on Fig. 3. On the other hand, as shown in Fig. 7(c) and (d), considering the phase volume fraction and characteristic length scale in objective function f_2 gets rid of the degeneracy issue and direct the BBO method toward a unique point in the composition-process space, yielding the target microstructure. In conclusion, it seems the integration of the volume fraction and characteristic length scale of phases in the objective function provides a more unique representation of microstructures in this phase-field case study, which highlights the importance of choosing proper microstructure descriptors in process design. However, it should be noted that objective function f_2 in this work is not necessarily the best for all microstructure-based inverse problems, and a different (more complicated) one can be defined using more microstructural features or QoIs if required.

As an example, the probabilistic GP models obtained during the entire BBO process with the second objective function and the batch size 100 are shown in Fig. 8. The starting GP model in Fig. 8(a) with three initial points are followed by the GP models after the first and second BBO iterations in Fig. 8(b) and (c). Blue and green surfaces in each

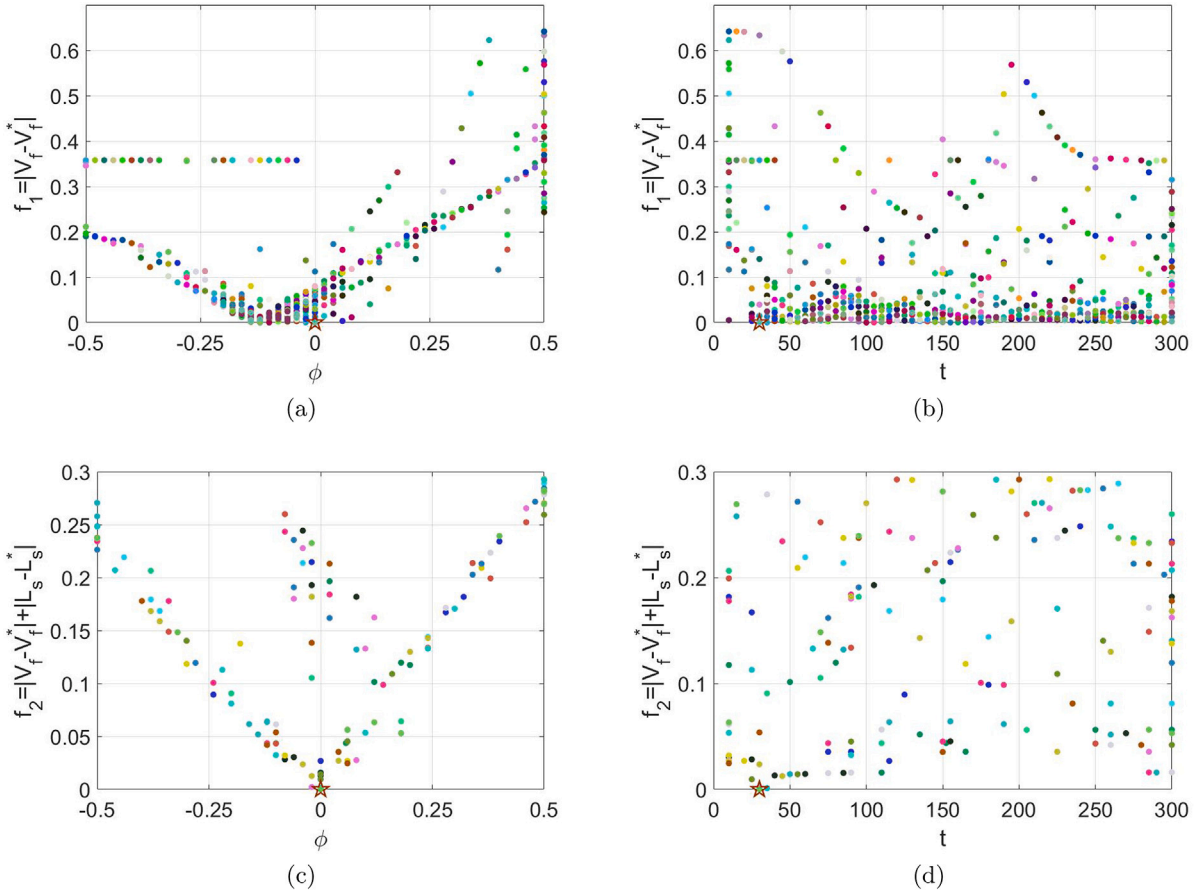

Fig. 7. The values of objective function in terms of field order parameter and time at the three initial points and multiple ten points ($K = 10$) suggested during the BBO iterations for the ((a), (b)) first and ((c), (d)) second objective function case. Colors indicate the points initially selected or suggested in each BBO iteration for each objective function case and the red pentagrams show the target field order parameter and time.

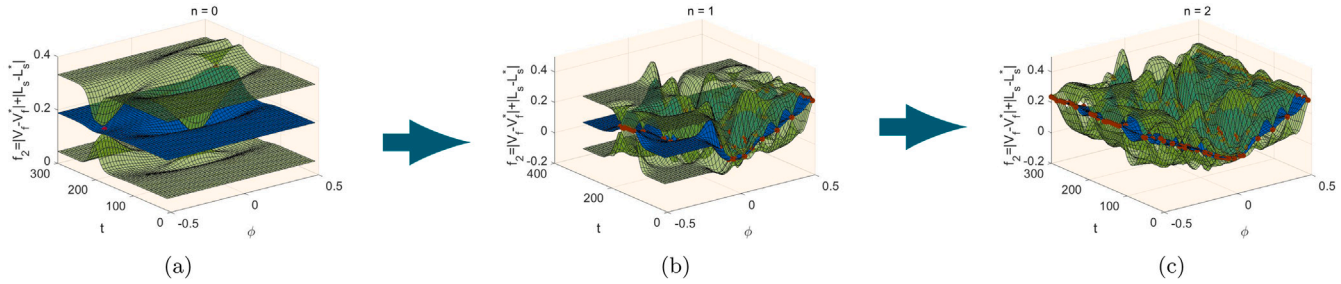

Fig. 8. Probabilistic GP models fitted during the BBO process with the second objective function and the batch size 100 on (a) the initial data points, (b) the initial data points plus the evaluated data points after the first optimization iteration, and (c) the initial data points plus the evaluated data points after the first and second optimization iterations. In these plots, blue and green surfaces correspond to the predictive mean values and 95% credible interval boundaries at any given point in the input space, respectively.

plot also indicate the predictive mean values and 95% credible interval boundaries at any given point in the input space, respectively. Again, a relatively steep gradient toward the optimum solution is observed in the GP predictive mean surface at the end of the optimization process, as shown in Fig. 8(c).

3.2.2. Computational microstructure-based process design using the high-fidelity model

In this case study, the design input space includes AlN composition (c_{AlN}) from 0.24 to 0.7 incremented by 0.02, ZrN composition (c_{ZrN}) from 0.02 to 0.08 incremented by 0.005, and time (t) from 10 to 50 incremented by 2, while TiN composition (c_{TiN}) is ignored due to its dependency on the other two elemental compositions. The output space is defined based on the second objective function in the previous study (f_2) as it was shown that it better represents the microstructures and leads to the optimum solution in less iterations. Similarly, the microstructural descriptors of two phases are taken into account here as the third phase descriptors are dependent on the descriptors of the other two phases. However, this objective function weights both microstructure descriptors equally, which is not necessarily the most optimal metric for the optimization processes. Therefore, a dynamic objective function (f_3) is defined in terms of the weights of the microstructure descriptors. In other words, dynamic weighting is performed by considering w and $1 - w$ as the coefficients of the characteristic length scale and volume fraction of phases in the objective function, respectively, i.e., $f_3 = (1 - w) \times |V_f - V_f^*| + w \times |L_s - L_s^*|$. Here, w is a hyperparameter that is sampled from 0 to 1 incremented by 0.05 besides the design input variables during the BBO process.

To evaluate the efficiency of this dynamic objective function compared to the equally weighted objective function in the BBO approach, a synthetic microstructure generated by the input values of $(c^*_{\mathrm{AlN}}, c^*_{\mathrm{ZrN}}, t^*) = (0.36, 0.05, 38)$ is considered as the design target and this structure is shown in Fig. 1(b). For both objective function cases, the BBO starts by fitting the GP models over the objective-function values obtained by the phase-field simulations at the same 50 initial points generated through Latin Hypercube Sampling (LHS) in the 3-D input space. The weight hyperparameter in the case of BBO with f_3 is considered to be 0.5 for all the 50 points in order to initialize both optimizations with exactly the same conditions. The optimization processes continue with batch size 50 in an iterative way until a microstructure associated with the target input point or its neighboring points in the 3-D space is achieved. A less strict stopping criterion is selected in this study to reduce the required number of optimization iterations since the applied high-fidelity phase-field model is computationally expensive to run. While only 1000 length-scale hyperparameters were sampled in the case of the low-fidelity model, 10000 length-scale hyperparameters are sampled here for the BBOs on the high-fidelity model due to the larger hyperparameter space in this study.

In Fig. 9, a 2-D t-distributed stochastic neighbor embedding (t-SNE) plot is shown for the 3-D design input space in yellow and samples suggested during the BBO iterations with objective function f_2 in different colors. In this plot, the target input point is also marked by a red pentagram. It is worth noting that t-SNE is a nonlinear, unsupervised ML technique, mainly used to visualize high-dimensional data in 2-D or 3-D space based on the pairwise distances between the data points in high-dimensions. The same t-SNE representation of the 3-D design input space is shown in Fig. 10(a) with samples associated with the BBO iterations with objective function f_3. Fig. 10(b) represents the same samples but mapped from a 4D space by considering their weight hyperparameter values. In this plot, there are 21 target input points corresponding to the target input values for compositions and time with 21 discretized values for the weight hyperparameter.

As shown in Figs. 9 and 10, it takes 7 and 3 iterations or $7 \times 50 = 350$ and $3 \times 50 = 150$ phase-field evaluations in addition to the 50 initial evaluations to achieve the optimal result. The number of evaluations until the optimal results implies the BBO with dynamic objective function f_3 is almost twice as efficient as the one with an objective function that equally weights the microstructure descriptors in this case study. The optimal microstructures obtained from these two optimizations have been shown in Fig. 11. As observed in this figure, these microstructures are very close to the target microstructure, indicating no specific changes in morphologies and the values of microstructure descriptors for the neighboring points in the design input space. Fig. 11 also shows that the target microstructure (part a) has closer descriptor values to the optimal microstructure obtained from optimization with objective function f_3 (part c) than the optimal microstructure obtained from optimization with objective function f_2 (part b) since the first two neighboring points are spatially closer.

Recently, the idea of using inexpensive ML models, such as deep neural networks (DNNs), as emulators of microstructure evolution simulations has become a hot topic in accelerated microstructure-based process design [33,54–56]. However, it is suspected that the number of phase-field evaluations required to train such ML models is higher than the evaluations suggested by the BBO approach. The reason is that the surrogate ML models can closely emulate the ground truth model only if trained with a relatively high number of data points throughout the entire input space, while the BBO technique guides the spatial search through the regions in the input space that are likely to deliver optimal results. Moreover, to the best of our knowledge, such ML models have yet to be applied to complex multi-physics models, where the model training becomes more challenging. This discussion indicates potentially higher power and efficiency of the applied BBO approach with the dynamic objective function than the competing cutting-edge approaches for materials process design based on the microstructure of interest. However, the claim above needs to be investigated practically by comparing the BBO approach with the other fast-acting design approaches in different materials design problems.

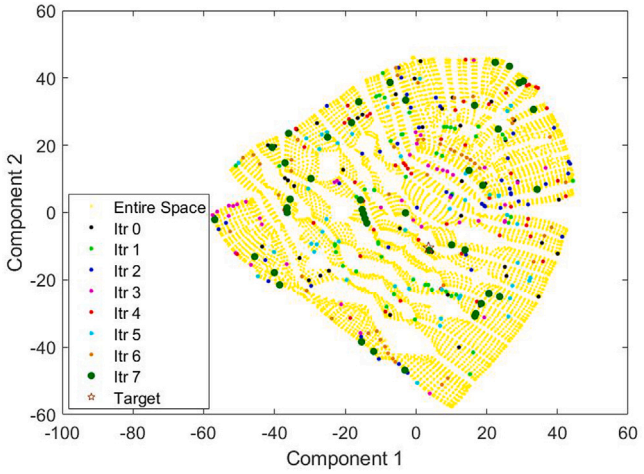

Fig. 9. 2-D t-SNE plot for the 3-D design input space and samples suggested during the BBO iterations with batch size 50 and objective function f_2 in different colors (Itr 0 refers to the initial samples generated by LHS). The red pentagram shows the target input point. It takes 7 BBO iterations to reach the target using objective function f_2.

4. Conclusions

In this work, a multi-scale search BBO method is proposed to address one of the most important challenges and concerns in the context of the ICME framework, i.e., the slow rate of materials design/discovery due to typically large (high-dimensional) composition–process spaces and/or super-expensive ground truth models and experiments. The proposed BBO method provides a parallel efficient search scheme in the materials composition–process spaces in order to tackle this problem. It was shown that the parallelization capability makes this method much more efficient than the commonly used standard BO method, where the one-at-a-time iterative selection of points restricts its application in most materials design problems. Also, the application of this BBO approach in the microstructure-based process design using two phase-field models—that is one of the most challenging inverse problems along the materials PSPP chain—indicated the importance of the proper definition of objective functions in such design problems. In the first example with the low-fidelity model for a binary spinodal decomposition, it is shown that the phase volume fractions are not enough to uniquely represent the spinodal microstructures, and their combination with the characteristic length scales is required to have a better objective function in BBO. In the second example with the high-fidelity model for a ternary spinodal decomposition, BBO with a dynamic objective function—where the weights for the microstructure descriptors, i.e., phase volume fractions and characteristic length scales, are considered random variables (or hyperparameters) and their values are selected beside the model input variables during the BBO process—is more efficient than the BBO with an objective function that equally weights these descriptors. In the end, this work attempts to highlight the value of the applied BBO approach as a promising pathway in accelerated materials design/discovery. However, this approach can be studied further using even more sophisticated multi-objective optimization algorithms in more complex design problems.

CRediT authorship contribution statement

P. Honarmandi: Conceptualization, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft, Writing – review & editing. **V. Attari:** Software, Visualization, Writing – original draft, Writing – review & editing. **R. Arroyave:** Funding acquisition, Project administration, Supervision, Writing – review & editing.

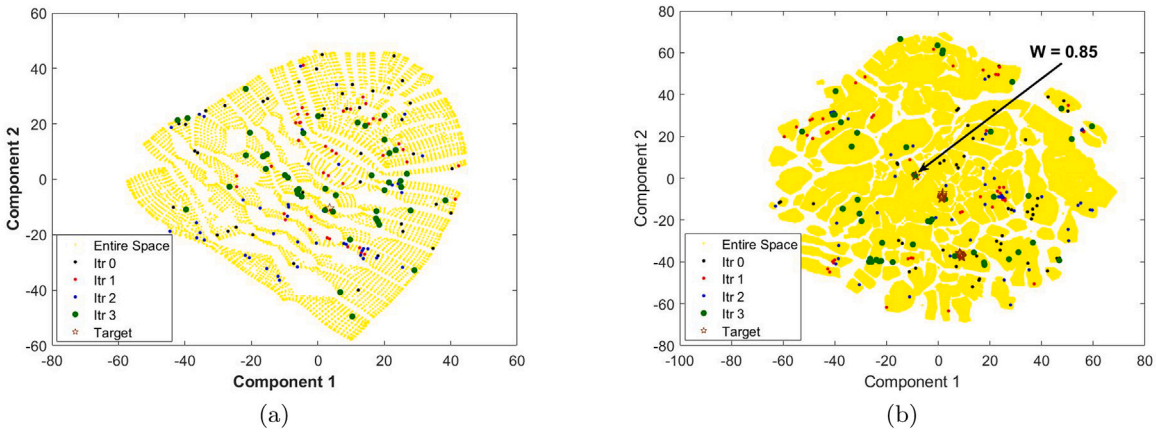

Fig. 10. 2-D t-SNE plots for (a) the 3-D design input space and (b) the 4-D space consisting of the 3-D design input space and the fourth dimension of weight hyperparameter, besides the samples suggested during the BBO iterations with batch size 50 and objective function f_3 in different colors (Itr 0 refers to the initial samples generated by LHS). The red pentagrams show the target input points. It takes 3 BBO iterations to reach the target using objective function f_3, which is less than half of the number of iterations obtained using objective function f_2.

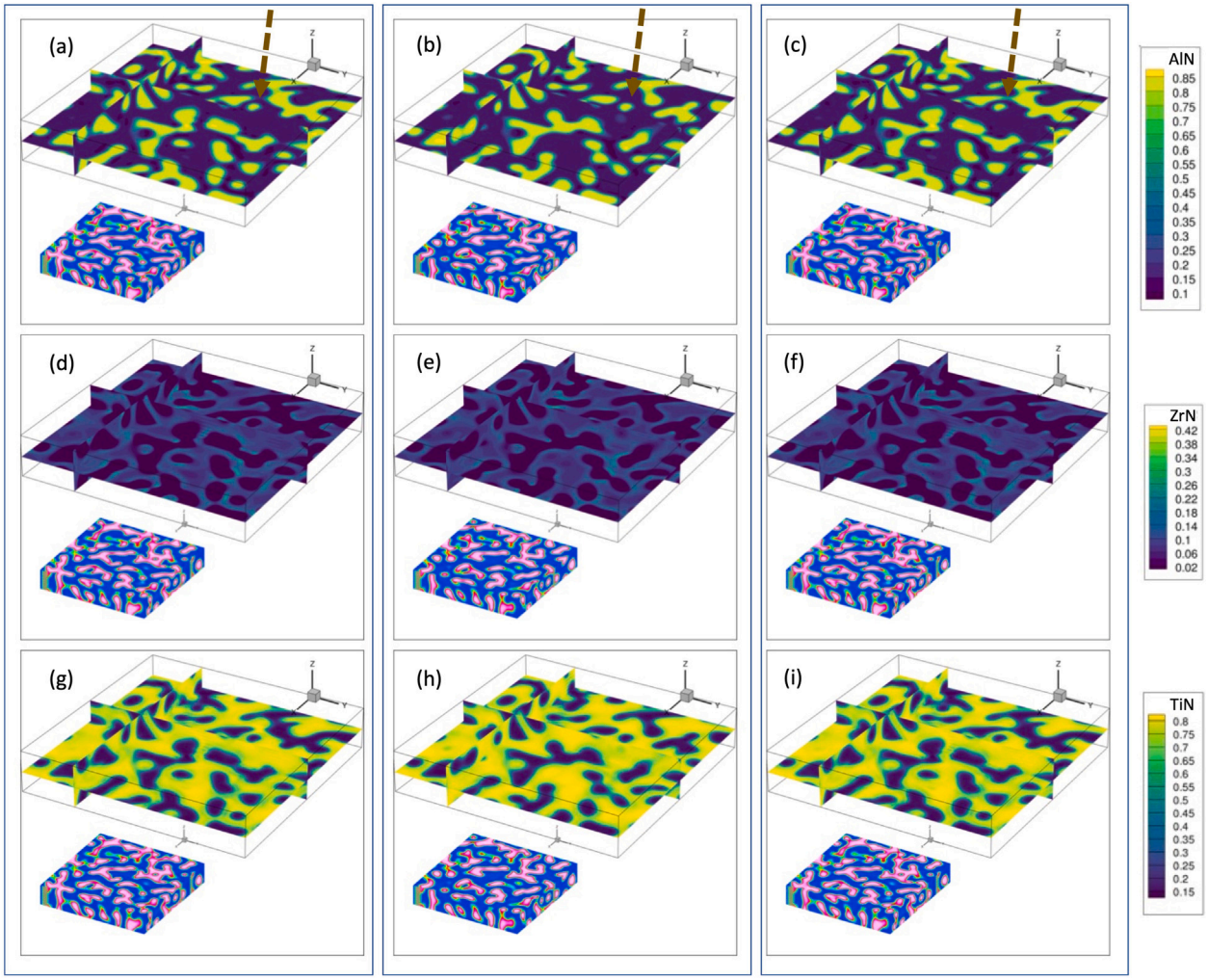

Fig. 11. 3D microstructures with cross section slices from three different planes obtained using high fidelity phase-field model for the (a, d, g) target microstructure - {input vector: $(c^*_{\mathrm{AlN}}, c^*_{\mathrm{ZrN}}, t^*) = (0.36, 0.05, 38)$, output vector (descriptors): $(V^*_{f-\mathrm{AlN}}, V^*_{f-\mathrm{ZrN}}, L^*_{s-\mathrm{AlN}}, L^*_{s-\mathrm{ZrN}}) = (0.355, 0.007, 2.116 \times 10^{-8}, 3.937 \times 10^{-10})$}, (b, e, h) optimal microstructure obtained from optimization with objective function f_2 - {input vector: $(c_{\mathrm{AlN}}, c_{\mathrm{ZrN}}, t) = (0.34, 0.05, 36)$, output vector (descriptors): $(V_{f-\mathrm{AlN}}, V_{f-\mathrm{ZrN}}, L_{s-\mathrm{AlN}}, L_{s-\mathrm{ZrN}}) = (0.324, 0.005, 1.577 \times 10^{-8}, 3.035 \times 10^{-10})$}, and (c, f, i) optimal microstructure obtained from optimization with objective function f_3 - {input vector: $(c_{\mathrm{AlN}}, c_{\mathrm{ZrN}}, t) = (0.36, 0.05, 36)$, output vector (descriptors): $(V_{f-\mathrm{AlN}}, V_{f-\mathrm{ZrN}}, L_{s-\mathrm{AlN}}, L_{s-\mathrm{ZrN}}) = (0.355, 0.007, 2.106 \times 10^{-8}, 3.804 \times 10^{-10})$}. The quivers point out a difference in the three microstructures. Top row (a, b, c) shows the AlN composition map, middle row (d, e, f) shows the ZrN composition map, and (g, h, i) shows the TiN composition map. The provided colorbars correspond to the cross section plots in each row. For the interpretation of numbers in these colorbars refer to the phase diagram of pseudo-ternary $\mathrm{Ti}_{1-x-y}\mathrm{Al}_x\mathrm{Zr}_y\mathrm{N}$ alloy [36]. For the interpretation of colors in this figure refer to the online version of this document.

Declaration of competing interest

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

Data availability

All data that support the findings of this study are openly available upon reasonable request.

Acknowledgments

PH and RA acknowledge the support of NSF, United States through Grant No. 1849085. VA and RA acknowledge the support by the National Science Foundation, United States under NSF Grants No. CMMI-1462255, CMMI-1663130, and DMR-2001333, and Texas A&M University Data Science Institute, United States (TAMIDS). Portions of this research were conducted with the advanced computing resources provided by Texas A&M High Performance Research Computing.

References

[1] A. Tran, J.A. Mitchell, L.P. Swiler, T. Wildey, An active learning high-throughput microstructure calibration framework for solving inverse structure–process problems in materials informatics, Acta Mater. 194 (2020) 80–92.
[2] R. Dehghannasiri, D. Xue, P.V. Balachandran, M.R. Yousefi, L.A. Dalton, T. Lookman, E.R. Dougherty, Optimal experimental design for materials discovery, Comput. Mater. Sci. 129 (2017) 311–322.
[3] T. Lookman, P.V. Balachandran, D. Xue, R. Yuan, Active learning in materials science with emphasis on adaptive sampling using uncertainties for targeted design, NPJ Comput. Mater. 5 (1) (2019) 1–17.
[4] A. Talapatra, S. Boluki, T. Duong, X. Qian, E. Dougherty, R. Arróyave, Autonomous efficient experiment design for materials discovery with Bayesian model averaging, Phys. Rev. Mater. 2 (11) (2018) 113803.
[5] D.R. Jones, M. Schonlau, W.J. Welch, Efficient global optimization of expensive black-box functions, J. Global Optim. 13 (4) (1998) 455–492.
[6] Y. Zhang, D.W. Apley, W. Chen, Bayesian optimization for materials design with mixed quantitative and qualitative variables, Sci. Rep. 10 (1) (2020) 1–13.
[7] B. Shahriari, K. Swersky, Z. Wang, R.P. Adams, N. De Freitas, Taking the human out of the loop: A review of Bayesian optimization, Proc. IEEE 104 (1) (2015) 148–175.
[8] S. Greenhill, S. Rana, S. Gupta, P. Vellanki, S. Venkatesh, Bayesian optimization for adaptive experimental design: A review, IEEE Access 8 (2020) 13937–13948.
[9] Y. Tenne, C.-K. Goh, Computational Intelligence in Expensive Optimization Problems, Vol. 2, Springer Science & Business Media, 2010.
[10] J. Azimi, A. Jalali, X. Fern, Hybrid batch Bayesian optimization, 2012, arXiv preprint arXiv:1202.5597.
[11] V. Nguyen, S. Gupta, S. Rana, C. Li, S. Venkatesh, Distance Exploration for Scalable Batch Bayesian Optimization.
[12] J. González, Z. Dai, P. Hennig, N. Lawrence, Batch Bayesian optimization via local penalization, in: Artificial Intelligence and Statistics, PMLR, 2016, pp. 648–657.
[13] T.T. Joy, S. Rana, S. Gupta, S. Venkatesh, Batch Bayesian optimization using multi-scale search, Knowl.-Based Syst. 187 (2020) 104818.
[14] A. Seko, A. Togo, H. Hayashi, K. Tsuda, L. Chaput, I. Tanaka, Prediction of low-thermal-conductivity compounds with first-principles anharmonic lattice-dynamics calculations and Bayesian optimization, Phys. Rev. Lett. 115 (20) (2015) 205901.
[15] D. Sterling, T. Sterling, Y. Zhang, H. Chen, Welding parameter optimization based on Gaussian process regression Bayesian optimization algorithm, in: 2015 IEEE International Conference on Automation Science and Engineering, CASE, IEEE, 2015, pp. 1490–1496.
[16] B. Rouet-Leduc, K. Barros, T. Lookman, C.J. Humphreys, Optimisation of GaN LEDs and the reduction of efficiency droop using active machine learning, Sci. Rep. 6 (1) (2016) 1–6.
[17] T. Lookman, P.V. Balachandran, D. Xue, J. Hogden, J. Theiler, Statistical inference and adaptive design for materials discovery, Curr. Opin. Solid State Mater. Sci. 21 (3) (2017) 121–128.
[18] S. Ju, T. Shiga, L. Feng, Z. Hou, K. Tsuda, J. Shiomi, Designing nanostructures for phonon transport via Bayesian optimization, Phys. Rev. X 7 (2) (2017) 021024.
[19] C. Li, D.R. de Celis Leal, S. Rana, S. Gupta, A. Sutti, S. Greenhill, T. Slezak, M. Height, S. Venkatesh, Rapid Bayesian optimisation for synthesis of short polymer fiber materials, Sci. Rep. 7 (1) (2017) 1–10.
[20] C. Sharpe, C.C. Seepersad, S. Watts, D. Tortorelli, Design of mechanical metamaterials via constrained bayesian optimization, in: ASME 2018 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers Digital Collection, 2018.
[21] M. Yamawaki, M. Ohnishi, S. Ju, J. Shiomi, Multifunctional structural design of graphene thermoelectrics by Bayesian optimization, Sci. Adv. 4 (6) (2018) eaar4192.
[22] R. Yuan, Z. Liu, P.V. Balachandran, D. Xue, Y. Zhou, X. Ding, J. Sun, D. Xue, T. Lookman, Accelerated discovery of large electrostrains in BaTiO3-based piezoelectrics using active learning, Adv. Mater. 30 (7) (2018) 1702884.
[23] P.V. Balachandran, B. Kowalski, A. Sehirlioglu, T. Lookman, Experimental search for high-temperature ferroelectric perovskites guided by two-step machine learning, Nature Commun. 9 (1) (2018) 1–9.
[24] A. Talapatra, S. Boluki, P. Honarmandi, A. Solomou, G. Zhao, S.F. Ghoreishi, A. Molkeri, D. Allaire, A. Srivastava, X. Qian, et al., Experiment design frameworks for accelerated discovery of targeted materials across scales, Front. Mater. 6 (2019) 82.
[25] C. Wen, Y. Zhang, C. Wang, D. Xue, Y. Bai, S. Antonov, L. Dai, T. Lookman, Y. Su, Machine learning assisted design of high entropy alloys with desired property, Acta Mater. 170 (2019) 109–117.
[26] A. Sakurai, K. Yada, T. Simomura, S. Ju, M. Kashiwagi, H. Okada, T. Nagao, K. Tsuda, J. Shiomi, Ultranarrow-band wavelength-selective thermal emission with aperiodic multilayered metamaterials designed by Bayesian optimization, ACS Central Sci. 5 (2) (2019) 319–326.
[27] A. Iyer, Y. Zhang, A. Prasad, S. Tao, Y. Wang, L. Schadler, L.C. Brinson, W. Chen, Data-centric mixed-variable bayesian optimization for materials design, in: International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, vol. 59186, American Society of Mechanical Engineers, 2019, p. V02AT03A066.
[28] A. Gilad Kusne, H. Yu, C. Wu, H. Zhang, J. Hattrick-Simpers, B. DeCost, S. Sarker, C. Oses, C. Toher, S. Curtarolo, et al., On-the-fly closed-loop autonomous materials discovery via Bayesian active learning, 2020, arXiv E-Prints, arXiv–2006.
[29] M.J. Borden, C.V. Verhoosel, M.A. Scott, T.J. Hughes, C.M. Landis, A phase-field description of dynamic brittle fracture, Comput. Methods Appl. Mech. Engrg. 217 (2012) 77–95.
[30] P.-A. Geslin, I. McCue, B. Gaskey, J. Erlebacher, A. Karma, Topology-generating interfacial pattern formation during liquid metal dealloying, Nature Commun. 6 (1) (2015) 1–8.
[31] P. Colli, G. Gilardi, D. Hilhorst, On a Cahn-Hilliard type phase field system related to tumor growth, 2014, arXiv preprint arXiv:1401.5943.
[32] J.W. Cahn, Phase separation by spinodal decomposition in isotropic systems, J. Chem. Phys. 42 (1) (1965) 93–99.
[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 Mater. 183 (2020) 452–470.
[34] L.Q. Chen, J. Shen, Applications of semi-implicit Fourier-spectral method to phase field equations, Comput. Phys. Comm. 108 (2–3) (1998) 147–158.
[35] V. Attari, A. Cruzado, R. Arroyave, Exploration of the microstructure space in TiAlZrN ultra-hard nanostructured coatings, Acta Mater. 174 (2019) 459–476.
[36] J. Zhou, L. Zhang, L. Chen, Y. Du, Z.-K. Liu, A thermodynamic description of metastable c-TiAlZrN coatings with triple spinodally decomposed domains, J. Min. Metall. B 53 (2) (2017) 85–93.
[37] A. Wang, S. Shang, D. Zhao, J. Wang, L. Chen, Y. Du, Z.-K. Liu, T. Xu, S. Wang, Structural, phonon and thermodynamic properties of fcc-based metal nitrides from first-principles calculations, CALPHAD 37 (2012) 126–131.
[38] N. Moelans, B. Blanpain, P. Wollants, An introduction to phase-field modeling of microstructure evolution, CALPHAD 32 (2) (2008) 268–294.
[39] A.G. Khachaturyan, Theory of Structural Transformations in Solids, Courier Corporation, 2013.
[40] T. Mura, Micromechanics of Defects in Solids, Springer Science & Business Media, 2013.
[41] H. Moulinec, P. Suquet, A numerical method for computing the overall response of nonlinear composites with complex microstructure, Comput. Methods Appl. Mech. Engrg. 157 (1–2) (1998) 69–94.
[42] M.P. Gururajan, T.A. Abinandanan, Phase field study of precipitate rafting under a uniaxial stress, Acta Mater. 55 (15) (2007) 5015–5026.
[43] R.A. Lebensohn, Y. Liu, P.P. Castaneda, Macroscopic properties and field fluctuations in model power-law polycrystals: full-field solutions versus self-consistent estimates, 460, (2045) The Royal Society, 2004, pp. 1381–1405,
[44] M. Larmuseau, M. Sluydts, K. Theuwissen, L. Duprez, T. Dhaene, S. Cottenier, Compact representations of microstructure images using triplet networks, NPJ Comput. Mater. 6 (1) (2020) 1–11.
[45] R.M. Haralick, K. Shanmugam, I.H. Dinstein, Textural features for image classification, IEEE Trans. Syst. Man Cybern. (6) (1973) 610–621.
[46] B.L. DeCost, E.A. Holm, A computer vision approach for automated analysis and classification of microstructural image data, Comput. Mater. Sci. 110 (2015) 126–133.

[47] D.T. Fullwood, S.R. Niezgoda, S.R. Kalidindi, Microstructure reconstructions from 2-point statistics using phase-recovery algorithms, Acta Mater. 56 (5) (2008) 942–948.
[48] C. Kunselman, V. Attari, L. McClenny, U. Braga-Neto, R. Arroyave, Semi-supervised learning approaches to class assignment in ambiguous microstructures, Acta Mater. 188 (2020) 49–62.
[49] C. Kunselman, S. Sheikh, M. Mikkelsen, V. Attari, R. Arróyave, Microstructure classification in the unsupervised context, Acta Mater. 223 (2022) 117434.
[50] N. Otsu, A threshold selection method from gray-level histograms, IEEE Trans. Syst. Man Cybern. 9 (1) (1979) 62–66.
[51] H. Liu, J. Cai, Y.-S. Ong, Y. Wang, Understanding and comparing scalable Gaussian process regression for big data, Knowl.-Based Syst. 164 (2019) 324–335.
[52] H.-S. Park, C.-H. Jun, A simple and fast algorithm for K-medoids clustering, Expert Syst. Appl. 36 (2) (2009) 3336–3341.
[53] S.C. Hardy, P.W. Voorhees, Ostwald ripening in a system with a high volume fraction of coarsening phase, Metall. Trans. A 19 (11) (1988) 2713–2721.
[54] E. Herman, J.A. Stewart, R. Dingreville, A data-driven surrogate model to rapidly predict microstructure morphology during physical vapor deposition, Appl. Math. Model. 88 (2020) 589–603.
[55] J. Yeom, T. Stan, S. Hong, P.W. Voorhees, Segmentation of experimental datasets via convolutional neural networks trained on phase field simulations, Acta Mater. (2021) 116990.
[56] V. Attari, R. Arroyave, Machine learning-assisted high-throughput exploration of interface energy space in multi-phase-field model with CALPHAD potential, Mater. Theory 6 (2022) 5, http://dx.doi.org/10.1186/s41313-021-00038-0.