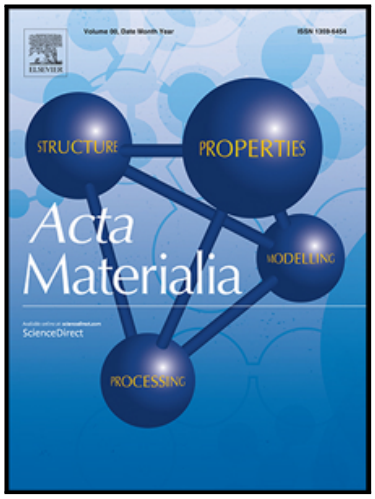
Journal Pre-proof



Efficiently exploiting process-structure-property relationships in material design by multi-information source fusion

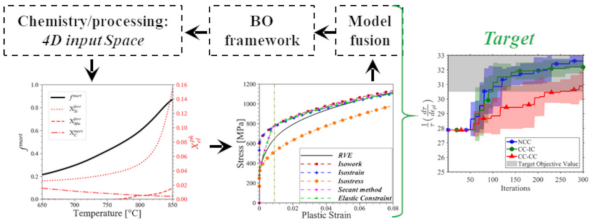
Danial Khatamsaz, Abhilash Molkeri, Richard Couperthwaite, Jaylen James, Raymundo Arróyave, Douglas Allaire, Ankit Srivastava

PII: S1359-6454(20)31056-9
DOI: https://doi.org/10.1016/j.actamat.2020.116619
Reference: AM 116619

To appear in: *Acta Materialia*

Received date: 15 August 2020
Revised date: 7 December 2020
Accepted date: 29 December 2020

Please cite this article as: Danial Khatamsaz, Abhilash Molkeri, Richard Couperthwaite, Jaylen James, Raymundo Arróyave, Douglas Allaire, Ankit Srivastava, Efficiently exploiting process-structure-property relationships in material design by multi-information source fusion, *Acta Materialia* (2020), doi: https://doi.org/10.1016/j.actamat.2020.116619



Efficiently exploiting process-structure-property relationships in material design by multi-information source fusion

Danial Khatamsaz[a], Abhilash Molkeri[b], Richard Couperthwaite[b], Jaylen James[b], Raymundo Arróyave[a,b,c], Douglas Allaire[a], Ankit Srivastava[b,*]

[a] *J. Mike Walker '66 Department of Mechanical Engineering, Texas A&M University, College Station, TX, USA*

[b] *Department of Materials Science and Engineering, Texas A&M University, College Station, TX, USA*

[c] *Department of Industrial and Systems Engineering, Texas A&M University, College Station, TX, USA*

Abstract

Materials design calls for the (inverse) exploitation of Process-Structure-Property (PSP) relationships to produce materials with targeted properties. Unfortunately, most materials design frameworks are not optimal, given resource constraints. Bayesian Optimization (BO)-based frameworks are increasingly used in materials design as they balance the exploration and exploitation of design spaces. Most BO-based frameworks assume that the design space can be queried by a single information source (e.g. experiment or simulation). Recently, we demonstrated microstructure-sensitive design of alloys with a BO framework capable of exploiting multiple information sources. While promising, the previous framework is limited as it assumes that the optimal microstructure is always feasible and considers microstructural features as the design space. Herein, we sidestep this unwarranted assumption and instead consider that chemistry and processing conditions constitute the design space amenable to optimization. We demonstrate the efficacy of our expanded framework by optimizing the mechanical performance of a ferritic/martensitic dual-phase material by adjusting composition/processing parameters. The framework uses thermodynamic results to predict microstructural attributes which are then used to predict the mechanical properties using a variety of micromechanical models and a microstructure-based finite element model. The final stage involves implementing model reification and information fusion, and a knowledge-gradient acquisition function to determine the next best design point and information sources to query. A detailed discussion of the various components and demonstration of how the framework can be implemented under three sets of cost-based constraints is presented.

Keywords: Computational thermodynamics; Micromechanical modeling; Two-phase materials; Mechanical properties; Bayesian optimization

* *Email address:* ankit.sri@tamu.edu (Ankit Srivastava)

1. Introduction

Materials design through ICME approaches [1] calls for the integration of physics (or data [2, 3])-based models and experiments in order to establish quantitative Process-Structure-Property (PSP) relationships [4], which can then be exploited in order to elucidate the mechanisms by which the (multi-scale, hierarchical) microstructure of a material responds to external stimuli (i.e. its property), or the mechanisms by which processing/synthesis conditions alters the way materials are (hierarchically) organized [5]. By 'inverting' such relationships, it is then possible to identify the combinations of chemistry and processing *necessary* to produce (multi-scale) microstructures that meet target measures of performance. [6].

Unfortunately, fully integrated ICME frameworks with quantitative predictive accuracy still remain out of reach, primarily due to the complex, highly coupled, multi-scale nature of linkages along the PSP chain. This complexity makes it very difficult to computationally emulate such PSP chains. Even if one ignores such challenges and focuses instead on much simpler single-scale or effective models, the explicit integration of multiple tools within a single ICME framework constitutes a major challenge [7] that has remained mostly unresolved, although limited recent works [8, 9] have achieved some success in deploying fully integrated PSP model chains to carry out the design/optimization of materials. Major challenges arise when considering the considerable cost associated with querying PSP relationships. Recently, closed-loop Bayesian Optimization (BO) approaches [10], capable of efficiently balancing the exploration and exploitation of materials design spaces, have already been shown to be quite effective in materials optimization tasks under resource constraints [11].

A major limitation of most approaches to date is the fact that they tend to use a single model per linkage along PSP chains. This is an unnecessary restriction as often there are multiple computational models or information sources to choose from when trying to establish such quantitative linkages. A further limitation is the fact that most 'classical' ICME frameworks are not capable of directly integrating experimental information into the design/optimization loop, other than for calibration, validation or verification purposes. Potential valuable experimental information could instead be directly incorporated into the ICME framework in a co-design scheme but this remains largely unexplored. In fact, even state-of-the-art BO-based materials design [10, 11] tends to be limited to a single probe (experimental or computational) to query specific linkages of the PSP chain, although other engineering fields have developed sophisticated approaches for the integration of multiple information sources within optimization schemes [12].

In recent work [13, 14], we have developed and demonstrated a materials design framework that addresses most of the issues highlighted above. In order to avoid explicit I/O interfaces between different models connecting microstructure to mechanical behavior—for a demonstration problem involving microstructural design of a dual-phase material—each of the models, including the 'ground truth' (in that case a microstructure-based finite element model) were converted into a Gaussian Process (GP). Since all models used are attempting to describe the same underlying behavior (i.e. the connection between microstructural features and mechanical response), it is to be expected that they would exhibit some degree

of correlation among themselves and also with the 'ground truth'. By exploiting such statistical correlations through the so-called 'reification' process [15, 16], we generated fused models that were capable of maximizing agreement with available information about the response of the 'ground truth' model, while minimizing responses clearly at odds with observations. The fused model was then used to answer two questions: (i) where to sample next in the design space, and (ii) which information source should be used to query the design space. To answer those questions, we constructed experimental utility functions based on the Knowledge Gradient (KG) [17]. In ref. [13], we demonstrated how this multi-information fusion BO framework was more efficient than state-of-the-art BO approaches that used a single (ground truth) information source. In ref. [14], we further demonstrated that explicitly accounting for the cost of individual information sources was essential if there are hard constraints imposed on the budget allocated to carry out optimization.

While promising, the previous framework was admittedly limited in that the complete PSP chains were not exploited and only a microstructure sensitive design problem was addressed, assuming that the design space consisted of a universe of microstructures that could in principle be accessed through suitable processing-chemistry combinations. This is a strong and unwarranted assumption, as *optimal microstructures may not necessarily be feasible* (i.e. accessible through the available chemistry and/or processing controls). It is thus important to explore possible ways of extending our multi-information source fusion BO framework in such a way that the design space consists of degrees of freedom that are truly amenable to modification without concerns on whether they are feasible or not. Such a framework would then truly realize the ICME program which relies, as mentioned above, on the exploitation of PSP relationships in order to carry out the design of materials. Solving such a problem is not trivial as we would have to fuse information not only in parallel, as we have shown before (all models used connected the same microstructural input to the same target mechanical response), but also in series, as we would have to provide explicit connections between processing-chemistry prescriptions and properties/performance through intermediate microstructure information.

In this work, we expand on our prior multi-information source fusion BO framework to account for cases in which it is possible to establish, at least in principle, quantitative PSP relationships. As in our previous works [13, 14], we demonstrate the framework by attempting to optimize the performance of a dual-phase steel. In this case, however, the degrees of freedom, amenable for optimization, that we consider are chemistry and processing (in this simplified case, heat treatment temperature), rather than microstructure characteristics. In this way we avoid the limitations of our previous works (and that of others) as in this case the optimum microstructure is always feasible since, it is a consequence of chemistry and processing protocol. We represent the proposed framework in Fig. 1. This framework optimizes the normalized strain hardening rate $(1/\tau)\,(\mathrm{d}\tau/\mathrm{d}\epsilon_{pl})$ of a dual-phase (ferrite-martensite) steel by adjusting the composition and heat-treatment temperature of the material. The framework uses a surrogate model of the thermodynamic results to predict the phase volume fraction and composition in the material microstructure after single-stage heat-treatment (intercritical annealing followed by quenching). This data is then used to predict the mechanical properties of the dual-phase material using a variety of micromechan-

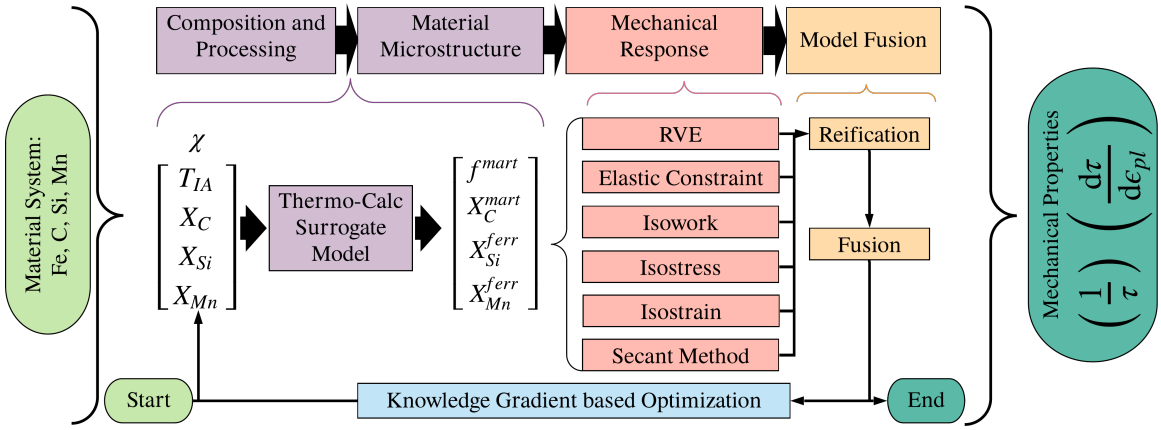


Figure 1: Schematic representation of the process-structure-property relationship used for the design of a dual-phase (ferrite-martensite) steel using our multi-information source fusion Bayesian optimization framework. Here, χ is the set of input variables with T_{IA} being the intercritical annealing temperature, X_C, X_{Si} and X_{Mn} being the carbon, silicon and manganese content, respectively, while the targeted output is the strength normalized strain hardening rate, $(1/\tau)\,(\mathrm{d}\tau/\mathrm{d}\epsilon_{pl})$.

ical models and a high through-put microstructure-based finite element model that utilizes a three-dimension representative volume element (RVE) of the material microstructure. After obtaining fused models through reification, the next point to query in the design space as well as the information source used to query it are determined using the KG acquisition function. What follows is a discussion of the various components of this framework and a demonstration of how the framework can be implemented under three separate sets of constraints, related to when queries to the 'ground truth' are made.

2. Methods

2.1. Connecting Chemistry and Processing to Microstructure

The current work considers a material system that contains C ($0.05-1$ wt%), Si ($0.1-2$ wt%), Mn ($0.15-3$ wt%) and Fe being the balance, and is subjected to a single-stage intercritical annealing heat-treatment at temperatures ranging from 650-850°C, followed by rapid quenching to produce a dual-phase (ferrite-martensite) microstructure. The prediction of the microstructure space is carried out through a surrogate model built from Thermo-Calc$^{\mathrm{TM}}$ predictions in the region of interest. This surrogate modeling approach was used to ensure that the calculations can be completed quickly, as well as to ensure that it would be possible to carry out optimization process on a computing resource without access to Thermo-Calc$^{\mathrm{TM}}$ license.

A detailed description on the construction of the surrogate model can be found in ref. [18]. Briefly, the Thermo-Calc$^{\mathrm{TM}}$ model was uniformly sampled in the design space to obtain the austenite volume fraction and composition. The volume fraction of martensite was obtained using the Koistinen-Marburger equation [19]. This data was fit with a Gaussian Process (GP) model. The composition and volume fraction of the ferrite phase are defined by a mass balance. This thermodynamic surrogate model was linked to the mechanical models

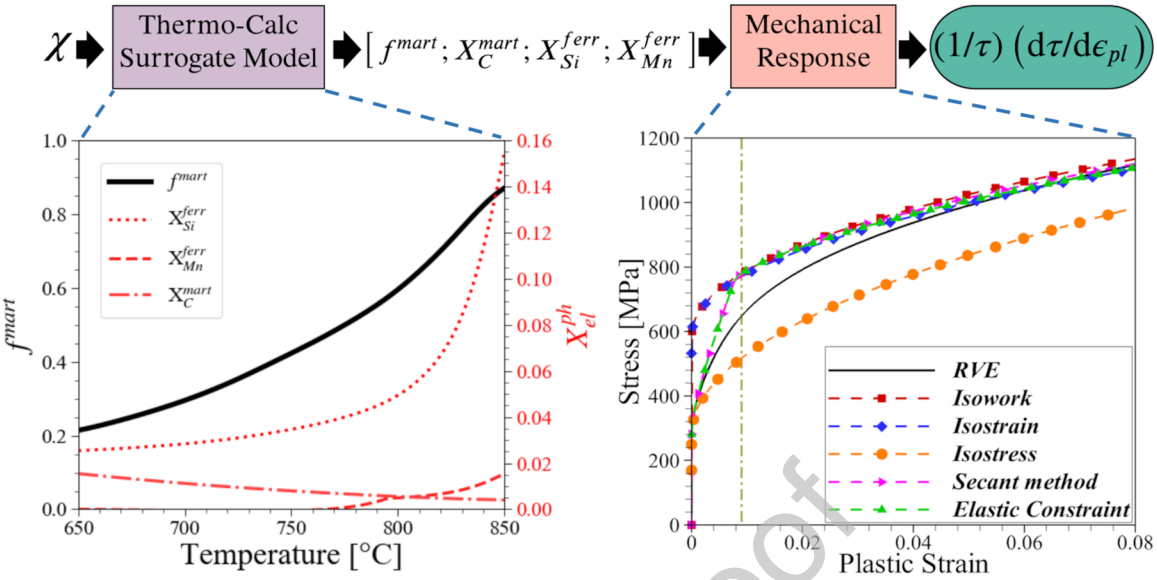


Figure 2: Schematic showing representative modeling results along the process-structure-property chain used herein for the design of a dual-phase (ferrite-martensite) steel. (Bottom left) Evolution of the martensite volume fraction, f^{mart}, and partitioning of alloying elements among the two constituent phases with temperature. (Bottom right) Evolution of stress with plastic strain in a dual-phase material as predicted by microstructure-based finite element model (RVE) and a variety of micromechanical models.

(described in the subsequent section) as shown in Fig. 2. As is often the case when using machine learning models, the accuracy of the GP surrogate is ultimately limited by the accuracy of the underlying ground truth as well as the ability of the model form itself to reproduce the data used to train it. In general, we have found [18] that GPs are able to model the outcome of phase equilibria calculations with reasonable accuracy.

2.2. Connecting Microstructure to Mechanical Response

In this work, we assume that we have at our disposal a set of models (or sources of information) of varying complexity (or computational cost) and fidelity that establish quantitative relationships between microstructure and properties.. The models specifically take as input the quantitative features of the dual-phase microstructure in order to predict its mechanical response. For the purpose of demonstration, we consider a microstructure-based finite element model as 'ground truth' and a variety of reduced-order and micromechanical models as alternate cheap sources of information. All the mechanical models predict the mechanical response of the material by explicitly incorporating the effect of both the chemistry and processing conditions—resulting in specific phase constitution—by utilizing the predictions of the Thermo-Calc^{TM} model, Fig. 2. Note that the stress-strain response of a composite dual-phase microstructure can in theory be accurately modeled using high-fidelity microstructure-based finite element calculations [20–22]. However, these calculations are computationally expensive and their direct use in an iterative material optimization scheme will be prohibitively time consuming. Reduced-order mechanical models [23–25] and more sophisticated micromechanical models [26] are computationally cheap, but this comes at a

price, since their many simplifications result in loss of fidelity with regards to the 'ground truth'—all the mechanical models are briefly described below.

2.2.1. Microstructure-based Finite Element Model (Ground truth)

The microstructure-based finite element calculations utilize a three-dimensional RVE of the dual-phase microstructure to compute the uniaxial tensile stress-strain response of the material [13, 14, 22]. To this end, we first generate a three-dimensional ferrite grain structure in the RVE using a simulated annealing process [27] with brick meshing. Next, a fixed number of martensite particles are nucleated randomly on the ferrite grain boundaries and then grown at a fast rate along the grain boundaries and a slow rate along other directions. The growth of the martensite particles is terminated when the overall volume fraction of the martensite reaches the prescribed target value. In the final RVE the crystallographic orientations of individual ferrite grains and that of martensite blocks are smeared out. This results in a composite dual-phase microstructure with two discretely modeled phases where each phase is modeled as an isotropic elastic-plastic material with Young's modulus, $E = 200$GPa, and Poisson's ration, $\nu = 0.3$, and constitutive relations (relating stress, τ, equivalent plastic strain, ϵ^{pl}, and composition, X_i^p, in weight fraction),

$$\tau = \tau_0^M + C_c(X_C^{mart})^{1/3} + K^M(\epsilon^{pl})^{n^M} \tag{1}$$

for the martensite phase with $\tau_0^M = 400$MPa, $C_c = 10^5$MPa, $K^M = 450$MPa and $n^M = 0.06$, and

$$\tau = \tau_0^F + C_{Si}(X_{Si}^{ferr})^{1/2} + C_{Mn}(X_{Mn}^{ferr})^{1/2} + K^F(\epsilon^{pl})^{n^F} \tag{2}$$

for the ferrite phase with $\tau_0^F = 200$MPa, $C_{Si} = 732$MPa, $C_{Mn} = 213$MPa, $K^F = 2200$MPa and $n^F = 0.5$. The constitutive relations in Eqs. (1) and (2), and the choice of the (representative) constitutive parameters are based on the assumptions that: (i) the strength of the martensite phase depends on its carbon content and the martensite phase doesn't exhibit significant strain-hardening, and (ii) the ferrite phase is softer than the martensite phase, exhibits significant strain-hardening and its strength depends on silicon and manganese content.

A typical finite element mesh of the RVE of the dual-phase microstructure employs 27,000 C3D8 brick elements from the ABAQUS/standard element library [28]. In the calculations, fully periodic boundary conditions are imposed on all six faces of the RVE. Next, to simulate uniaxial tensile deformation, the RVE is subjected to a monotonically increasing tensile strain along one of the axis, while maintaining zero resultant forces on the faces normal to the other two axes as in ref. [21]. Note that since the martensite particles are nucleated at random, *different realization of the same microstructure (i.e. the same overall phase constitution) in the RVE is different*. Thus, the mechanical response predicted by different realizations of the same microstructure in the RVE is also slightly different as demonstrated in ref. [14].

2.2.2. Reduced-order and micromechanical models (Information sources)

As in refs. [13, 29], low-fidelity models are used as information sources to link material microstructure to overall mechanical response. The reduced-order mechanical models used here are based on three assumptions on how the work, stress, and strain partition among the constituent phases: (i) an *isostrain model* is built on the assumption that both the constituent phases undergo the same amount of strain under deformation [23]; (ii) an *isostress model* is constructed following the assumption of equal partitioning of stresses among both the phases [24]; and (iii) an *isowork model*, that assumes that both the phases undergo the same amount of work of deformation upon the deformation of the dual-phase microstructure [25]. In addition, we employ two slightly more sophisticated micromechanical models linking material microstructure to overall mechanical response. These models are constructed based on the homogenization schemes proposed in ref. [26]. We refer to the first micromechanical model as the 'secant method', where the prediction of overall mechanical response is based on Hill's weakening constraint power in a plastically-deforming matrix. The second micromechanical model is referred to as the 'elastic constraint' model, which is based on Kröner's treatment of the matrix-inclusion system under elastic constraints [26]. All these reduced-order and micromechanical models use the same constitutive relations given in Eqs. (1) and (2) for the two constituent phases.

2.3. Material Design Framework

An optimization problem can be written as

$$\mathbf{x}^* = \arg\max_{\mathbf{x}\in\mathcal{X}} f(\mathbf{x}), \tag{3}$$

where f is the objective function and $\mathbf{x}^*$ is the optimal design vector in the feasible input space $\mathcal{X}$. Oftentimes, the analytical form of the objective function f is unknown or is very expensive to evaluate through either physical experiments or high fidelity simulations. Indeed, resource constraints—i.e. the total budget available to the design campaign—place strong limits on the number of times the objective function can be queried. In many cases, however, it is likely that there exist cheaper information sources that can potentially be used to *estimate* the response of the much more expensive information source, at varying degrees of fidelity and cost. Such information sources could be models constructed with simplifying assumptions or evaluated at much lower resolution than the 'ground truth'. In cases in which the 'ground truth' is an experiment, the cheap sources could be (physics-based or machine learning) models, or could be other experiments that are much faster/cheaper to carry out but that are correlated to the more expensive 'ground truth'.

Here, we propose a multi-fidelity approach to seek the solution to Eq. (3). Assuming that every source contains some useful information about the quantity of interest—different sources may approximate the 'ground truth' better than others in some regions of the design space—, we aim to efficiently fuse all these sources in order to estimate the objective quantity of interest as accurately as possible. Following our previous work [13, 29], we employ a multi-information source optimization framework. In this approach, the first step is to construct intermediate surrogates for each of the sources according to the prior knowledge

about the connection between the design space and their response. Here we use GPs as surrogates given their mathematical properties, including their ability to predict not only the mean value, but also the variance of the quantity of interest in the design space, and the straightforward manner in which the causal correlation between points in the design space can be modeled, among others [30]. Using a reification-based information fusion approach developed in ref. [15, 16], a fused model is built from all the information sources by exploiting the degree to which they are correlated with each other and with the 'ground truth'. The fused model is then used to determine the next design vector and information source to query. Here, we use the Knowledge Gradient (KG) acquisition function as the utility function to rank the next best design point to explore, and the best information source to query it, taking into account the cost of evaluating said information source as well as the uncertainty associated to the model predictions. This policy balances the exploration of the input space and exploitation of the current state of knowledge about the objective function, while also accounting for the impact of such exploration/exploitation on the available resources.

By fitting GPs to data from previous evaluations of the individual information sources, the prior GP distribution for each information source is represented as

$$f_{GP,i}(\mathbf{x}) \sim \mathcal{GP}\left(m_i(\mathbf{x}), k_i(\mathbf{x}, \mathbf{x}')\right), \tag{4}$$

where $k_i(\mathbf{x}, \mathbf{x}')$ is a real-valued kernel function over the input space and $m_i(\mathbf{x})$ defines the mean function. Here, we employ the squared exponential covariance function—note that other kernels may be used depending on our knowledge of the properties of the design space—as the kernel function specified as

$$k_i(\mathbf{x}, \mathbf{x}') = \sigma_s^2 \exp\left(-\sum_{h=1}^{d} \frac{(x_h - x_h')^2}{2l_h^2}\right), \tag{5}$$

where variables d, σ_s^2, and l_h are the dimension of the input space, signal variance, and characteristic length-scale. The latter controls the degree to which two points in the input space, separated by a specific distance, are correlated. The behavior of the GP surrogate is controlled by these hyperparameters and it is necessary to tune them based on the available data. Here, we optimize the hyperparameters by maximizing the log marginal likelihood.

Assuming N_i evaluations of information source i are available and indicated by $\{\mathbf{X}_{N_i}, \mathbf{y}_{N_i}\}$, where $\mathbf{X}_{N_i} = (\mathbf{x}_{1,i}, \ldots, \mathbf{x}_{N_i,i})$ is the N_i input samples to information source i and $\mathbf{y}_{N_i} = (f_i(\mathbf{x}_{1,i}), \ldots, f_i(\mathbf{x}_{N_i,i}))$ represents the outputs from information source i relatively, the posterior GP distribution of information source i at a design point $\mathbf{x}$ is given as

$$f_{GP,i}(\mathbf{x}) \mid \mathbf{X}_{N_i}, \mathbf{y}_{N_i} \sim \mathcal{N}\left(\mu_i(\mathbf{x}), \sigma_{GP,i}^2(\mathbf{x})\right), \tag{6}$$

where

$$\mu_i(\mathbf{x}) = K_i(\mathbf{X}_{N_i}, \mathbf{x})^T [K_i(\mathbf{X}_{N_i}, \mathbf{X}_{N_i}) + \sigma_{n,i}^2 I]^{-1} \mathbf{y}_{N_i}, \tag{7}$$

$$\sigma_{GP,i}^2(\mathbf{x}) = k_i(\mathbf{x}, \mathbf{x}) - K_i(\mathbf{X}_{N_i}, \mathbf{x})^T [K_i(\mathbf{X}_{N_i}, \mathbf{X}_{N_i}) + \sigma_{n,i}^2 I]^{-1} K_i(\mathbf{X}_{N_i}, \mathbf{x}), \tag{8}$$

and for information source i, $K_i(\mathbf{X}_{N_i}, \mathbf{X}_{N_i})$ is a $N_i \times N_i$ matrix with (m, n) entry as $k_i(\mathbf{x}_{m,i}, \mathbf{x}_{n,i})$, and $K_i(\mathbf{X}_{N_i}, \mathbf{x})$ is the $N_i \times 1$ vector whose m^{th} entry is $k_i(\mathbf{x}_{m,i}, \mathbf{x})$. Here, the

term $\sigma^2_{n,i}$ is used to model observation error for information source i found by experimental data.

We then quantify the discrepancy corresponding to each information source surrogate with respect to the 'ground truth' (i.e. the RVE model) as

$$\sigma^2_i(\mathbf{x}) = \sigma^2_{GP,i}(\mathbf{x}) + \sigma^2_{f,i}(\mathbf{x}), \tag{9}$$

where $\sigma^2_{f,i}(\mathbf{x})$ is the variance related to the fidelity of information source i, which is added to the uncertainty associated with the corresponding GP to find the total uncertainty of the information source surrogate i.

Since all information sources are estimating the same quantity of interest, they are expected to be correlated to the 'ground truth' to varying degrees. Every information source is assumed to have some useful information about the true objective function. In the proposed approach of refs. [13, 29], and contrary to conventional multi-fidelity approaches, it is not required to determine the hierarchy of fidelity according to agreement with the 'ground truth', rather *all information sources are brought together* into a single fused model. Here, we follow the same approach. Specifically, for fusion of normally distributed data—which is the case here since the information sources are represented by GPs—the method presented by ref. [31] is followed. This leads to fused mean and fused variance estimates at a given design point $\mathbf{x}$ as:

$$\mathbb{E}[\hat{f}(\mathbf{x})] = \frac{\mathbf{e}^{\mathrm{T}}\tilde{\Sigma}(\mathbf{x})^{-1}\boldsymbol{\mu}(\mathbf{x})}{\mathbf{e}^{\mathrm{T}}\tilde{\Sigma}(\mathbf{x})^{-1}\mathbf{e}}, \tag{10}$$

$$\mathrm{Var}\Big(\hat{f}(\mathbf{x})\Big) = \frac{1}{\mathbf{e}^{\mathrm{T}}\tilde{\Sigma}(\mathbf{x})^{-1}\mathbf{e}}, \tag{11}$$

where $\mathbf{e} = [1, \ldots, 1]^{\mathrm{T}}$, $\boldsymbol{\mu}(\mathbf{x}) = [\mu_1(\mathbf{x}), \ldots, \mu_S(\mathbf{x})]^{\mathrm{T}}$ are the mean values of S sources, and $\tilde{\Sigma}(\mathbf{x})^{-1}$ is the inverse of the covariance matrix between the information sources. For example, in the presence of two information sources, Eq. (10) is defined as

$$\mathbb{E}[\hat{f}(\mathbf{x})] = \frac{(\sigma_2^2 - \rho\sigma_1\sigma_2)\mu_1 + (\sigma_1^2 - \rho\sigma_1\sigma_2)\mu_2}{\sigma_1^2 + \sigma_2^2 - 2\rho\sigma_1\sigma_2}, \tag{12}$$

where σ_1^2 and σ_2^2 are the total variances of sources 1 and 2 respectively and the coefficient ρ shows the correlation between the information sources at a specific point $\mathbf{x}$. To estimate the correlations between the errors of two information sources i and j, we use reification [15, 16]. First, information source i is reified and the correlation coefficient is computed as

$$\rho_{ij}(\mathbf{x}) = \frac{\sigma_i^2(\mathbf{x})}{\sigma_i(\mathbf{x})\sigma_j(\mathbf{x})} = \frac{\sigma_i(\mathbf{x})}{\sqrt{(\mu_i(\mathbf{x}) - \mu_j(\mathbf{x}))^2 + \sigma_i^2(\mathbf{x})}}. \tag{13}$$

Then, information source j is reified to estimate $\rho_{ji}(\mathbf{x})$. Finally, the variance-weighted average of the two estimated correlation coefficients is computed to estimate the correlation between the models as

$$\bar{\rho}_{ij}(\mathbf{x}) = \frac{\sigma_j^2(\mathbf{x})}{\sigma_i^2(\mathbf{x}) + \sigma_j^2(\mathbf{x})}\rho_{ij}(\mathbf{x}) + \frac{\sigma_i^2(\mathbf{x})}{\sigma_i^2(\mathbf{x}) + \sigma_j^2(\mathbf{x})}\rho_{ji}(\mathbf{x}). \tag{14}$$

The average correlations are used in Eqs. (10) and (11) to estimate the fused mean and variance.

We use the fused means and variances to construct the fused model. Assuming N_f samples from the design space $\mathcal{X}$ are available, the fused mean vector and the diagonal matrix of fused variances are $\mu_{\text{Wink}}(\mathbf{x}_{1:N_f})$ and $\sum_{(\mathbf{x}_{1:N_f})} = \text{diag}\ (\sigma^2_{\text{Wink}}(\mathbf{x}_1), ..., \sigma^2_{\text{Wink}}(\mathbf{x}_n))$ respectively. The posterior predictive distribution of the fused model is specified as

$$\hat{f}^{\text{ fused}}(\mathbf{x}) \sim \mathcal{N}(\mu^{\text{fused}}(\mathbf{x}), \Sigma^{\text{fused}}(\mathbf{x})). \tag{15}$$

Next, the fused GP, which is considered as our best predictive model containing all current knowledge about the 'ground truth', is used to determine the next design point and information source to query. This is done by employing a utility function that considers the cost and the value of such a query relative to the optimization of the objective function in hand.

The value-gradient utility, which is a cost-aware KG utility, is used to evaluate different information sources to compare the amount of knowledge gained about the objective function's maximum and the cost to determine which model and design point adds the most value regarding Eq. (3). The utility function searches for immediate improvement to the knowledge state of the system in one step and at the same time looks for the expected improvement in two steps where it has the highest gradient in knowledge. Using the fused GP and its maximum mean function value, the immediate expected improvement can be defined.

At a design point $\mathbf{x}$, defining information sources by $(i_{1:N}, \mathbf{x}_{1:N}, y_{1:N})$ with design points and corresponding objective values for the first N queries and taking $\hat{f}$ as the posterior distribution of the fused model, the expected improvement is given as

$$\begin{aligned} EI(\mathbf{x}) &= \mathbb{E}\left[\max_{\mathbf{x}'\in\mathcal{X}} \mathbb{E}[\hat{f}(\mathbf{x}')|i_{1:N}, \mathbf{x}_{1:N}, \mathbf{x}_{N+1} = \mathbf{x}, y_{1:N}] - \max_{\mathbf{x}'\in\mathcal{X}} \mathbb{E}[\hat{f}(\mathbf{x}')|i_{1:N}, \mathbf{x}_{1:N}, y_{1:N}]\right] \\ &= \mathbb{E}\left[\max_{\mathbf{x}'\in\mathcal{X}} \mathbb{E}[\hat{f}(\mathbf{x}')|i_{1:N}, \mathbf{x}_{1:N}, \mathbf{x}_{N+1} = \mathbf{x}, y_{1:N}]\right] - \max_{\mathbf{x}'\in\mathcal{X}} \mathbb{E}[\hat{f}(\mathbf{x}')|i_{1:N}, \mathbf{x}_{1:N}, y_{1:N}], \end{aligned} \tag{16}$$

where the last expression can be removed from the expectation operator since it is known when conditioned on the first N queries. Next, we use KG policy [17, 32, 33] to maximize this expectation. Defining the value of being at the knowledge state H^N as $V^N(H^N) = \max_{\mathbf{x}\in\mathcal{X}} H^N$, where the knowledge state itself is presented by $H^N = \mathbb{E}[\hat{f}(\mathbf{x}')|i_{1:N}, \mathbf{x}_{1:N}, y_{1:N}]$, the KG as a measure of expected improvement is

$$\nu^{\text{KG}}(\mathbf{x}) = \mathbb{E}[V^{N+1}(H^{N+1}(\mathbf{x})) - V^N(H^N)|H^N], \tag{17}$$

where the design point chosen is the point that maximizes $\nu^{KG}(\mathbf{x})$. To consider the immediate and expected improvement, we use the value-gradient utility given as

$$U = \mu^*_{\text{fused}} + \max_{\mathbf{x}\in\mathcal{X}} \nu^{\text{KG}}(\mathbf{x}). \tag{18}$$

Considering Eq. (18), the value-gradient policy chooses the next point to query such that value-gradient utility is maximized.

Searching for the next point to query begins by generating Latin hypercube samples (LHS) from the input space as alternative points. When an alternative point is queried from the $\mathrm{GP_i}$, based on Eq. (4), the output is presented as a normal distribution with a mean and variance. Then, N_q independent samples are drawn out of this normal distribution

$$f_i^q(\mathbf{x}) \sim \mathcal{N}(\mu_i(\mathbf{x}), \sigma^2_{\mathrm{GP},i}(\mathbf{x})) \quad \text{for } q = 1, \ldots, N_q, \tag{19}$$

with $f_i^q(\mathbf{x})$ as the sample q for alternate point $\mathbf{x}$ and information source i. To calculate the value-gradient utility for each alternate point, we augment $\mathrm{GP_i}$ with a sample drawn from $\mathrm{GP_i}$ to build the $\mathrm{GP_{temp,i}}$. This approach assumes that the alternate point and the predicted sample value are part of the system's knowledge and a new fused model is created. Using Eq. (18), the value-gradient utility is measured. The process is repeated for all samples drawn from the $\mathrm{GP_i}$ for the same alternative point. For every alternative point and information source, the expected value-gradient utility is calculated by

$$EU_{\mathbf{x},i} = \frac{1}{N_q} \sum_{q=1}^{N_q} U^q_{\mathbf{x},i}. \tag{20}$$

To consider the cost of querying information sources, the expected utility function can be found per unit cost. Finally, the alternative point with the highest expected value gradient utility per unit cost and the information source to which that alternative point was added is chosen for the next query. This process is repeated and the fused GP is updated until the budget is exhausted or convergence requirements are fulfilled. Then, by searching the fused GP's mean function, the optimum solution of Eq. (3) is estimated.

2.4. Case studies

There are many ways to approach an optimization problem based on how the search policy is employed and how the problem is constrained. In our work, three approaches are considered. The first two approaches are cost-controlled, but with different triggers to call the RVE model (or ground truth) and which components of the calculation are considered in the cost. The third approach has no cost consideration in executing different information sources. For every particular case-study, five realizations of the entire process are obtained and the results presented are the averaged values over these five realizations. The three different approaches are described in the following subsections.

2.4.1. Cost Consideration - Iteration Controlled 'Ground Truth' Query (CC-IC)

When querying information sources while limited resources are available for the design process, it is important to compute the value that each query adds to the system, while also considering the resources needed to execute the query. For our cost consideration - iteration controlled (CC-IC) approach, at each step, we select an alternative point and an information source to query that adds the highest value per unit cost to the system's state of knowledge. Considering the cost of queries influence the decision making process with exploitation of the design space being favored over its exploration. The decision-maker in

this case acts conservatively in order to avoid the risk of wasting the (constrained) resources on highly uncertain queries unless it can expect a very good 'return on the investment' of querying that source. To update the model discrepancies, the 'ground truth'—i.e. the RVE—is queried once after every ten evaluations of the information sources. In Table 1, the costs assigned to each information source are shown based on their computation time [14].

Table 1: Computational cost of various information sources.

Information source	Cost (Seconds)	Normalized cost
Isostrain	2.3×10^{-4}	1
Isostress	1.0×10^{-3}	4.4
Isowork	4.7×10^{-1}	2.0×10^{3}
Secant method	8.4×10^{1}	3.7×10^{5}
Elastic constraint	3.6×10^{1}	1.6×10^{5}
RVE	7.2×10^{3}	3.1×10^{7}

2.4.2. Cost Consideration - Cost Controlled 'Ground Truth' Query (CC-CC)

We also study a case in which the cost of the entire process is considered. This is referred to here as the cost consideration - cost controlled (CC-CC) approach. While querying information sources is costly, so too are constructing GP surrogates and calculating the fused model and utility function, especially as the number of data points increases over time. We call this the modeling cost and measure it by calculating the time required for (i) generating alternative points and samples, (ii) building temporary GPs, (iii) performing the reification process, information fusion, and building the fused GP, (iv) evaluating the utility function, and (v) deciding on the next point to query. The measurement of the modeling cost, including the computational cost of the individual sources and the cost of generating points to sample and selecting the sources to query, is completed on each iteration and subtracted from the budget.

Additionally, we define a cost-controlled criterion to query the 'ground truth' model. Basically, instead of querying the 'ground truth' after a specific number of queries of the (cheaper) information sources, a query is made when a specific amount of the budget is spent. This strategy allows the decision-maker to use the cheaper information sources before querying the expensive 'ground truth' model, as long as it is finding value in querying the cheap information sources. This policy is in a sense the closest analogue to a realistic materials design campaign, where the budget available greatly determines the querying (i.e. experimentation or simulation) protocol.

Considering the modeling cost *prevents the framework from querying a very large number of points from the cheaper information sources before querying the 'ground truth'*. For instance, thousands of queries are needed to pass any defined amount of budget if the isostrain information source is chosen each time. This has the potential of significantly increasing the overall cost. By considering the modeling cost, fewer calls to the cheaper models are made, which helps to moderate the increase in the modeling cost.

2.4.3. No Cost Consideration (NCC)

When the cost of evaluating information sources is not considered as part of the decision making process, the system always chooses the alternative point and the information source to query with the highest expected value gradient utility, no matter how expensive the query is. In other words, there is no constraint on the decision making process concerning costs. In this approach, which we refer to as no cost consideration (NCC), the 'ground truth' is queried after every ten information source evaluations in the step to update the model discrepancy terms. This strategy leads to increased exploration of the design space, thus adding the maximum knowledge from the evaluation to the system after each update. Although this approach is not considered practical in real-world optimization problems, it is used as a comparison with the cost-controlled approaches in the current work.

3. Results

We now present the results of applying our material design framework to optimize the normalized strain hardening rate, $(1/\tau)(d\tau/d\epsilon^{pl})$ at a plastic strain level of 0.9%, of a dual-phase (ferrite-martensite) steel by adjusting the composition and heat-treatment temperature of the material. The design problem follows the three decision-making policies described in Section 2.4. To compare these decision-making policies, we consider how quickly each case reaches what is considered as the maximum mechanical property region of the design space. In our previous work ref. [14], for a fixed material chemistry, the maximum objective value was found to be 30.5, with a certain level of uncertainty due to stochastic nature of the process—related to the different test points arising from Latin hypercube sampling. In the current work, we seek to find an objective value as good as the best in ref. [14] or better. Accordingly, a normalized strain hardening rate ($[1/\tau][\mathrm{d}\sigma/\mathrm{d}\epsilon]$) greater than 30.5 is considered as the target objective value to reach and this is shown by the shaded area in Figs. 3 and 4.

To quantify the uncertainty in the current design process, five realizations of the design process were generated. Each realization utilized the same initial data and was allowed to run to completion. Through this approach we were able to quantify the combined effect of several sources of uncertainty in the optimization process. The first is that the RVE calculations rely on a stochastic approach to generate the dual-phase microstructure. As a result, the prediction from RVE calculation can be different even for the exact same inputs that include volume fraction and composition of the constituent phases. The next source of uncertainty that needs to be accounted for is the determination of the points at which the KG is evaluated. This is done by Latin Hypercube sampling and as a result will vary for each realization of the framework. The final source of uncertainty arises while calculating the KG, since a set of alternate points are sampled randomly from the normal distribution defined by the GP at the chosen test point. The KG is calculated for all these alternate points and then averaged. This average KG value is then used when calculating the test point with the maximum KG. Given the considerable computational cost of the scheme it is not feasible, at this time, to deconvolute these uncertainty sources, and we are thus only

estimating the total uncertainty and depicting it as shaded uncertainty bounds in Figs. 3 and 4.

As a preliminary step to set up the design problem, it is important to study how the prior belief about the 'ground truth' result provided to the system can affect the final results. This prior information is defined as the number of initial RVE evaluations (or 'ground truth' information) used to train the intermediate GP for the RVE information source. The goal is to use an initial training set size that sufficiently adds information while avoiding any extra expense on generating unnecessarily large initial training data sets.

The results of the parametric study conducted to address this question are presented in Table 2. Specifically, the effect of different initial training set sizes of the RVE on iterations and cost required to reach the initial 'guess' maximum objective value is investigated. The parametric study was carried out using the NCC decision-making policy, where during the optimization, the RVE is queried every ten iterations or queries to the cheaper information sources. The cost shown is divided into cost of generating the initial training set and the cost incurred during the optimization process. The final objective value reached for all the cases is the maximum normalized strain hardening rate found after reaching 300 iterations. We observe that from an initial training set size of 20 to 40, the iterations needed to reach the 'guess' maximum objective value decreases by half. However, as we increase the size of the initial training set from 40 to 70, we observe no significant change in the number of iterations required to reach the objective value and *it is thus a waste of resources to generate training sets greater than* 40. Similarly, we do not see any major improvement in the final objective value reached at the end of the optimization by increasing the initial training set from 40 to 70. A similar effect is observed when a training set size of 100 is considered. Therefore, to balance the cost of training and optimization, for all calculations presented hereafter, we use an initial training set size of 40 RVE points.

Table 2: Comparing the number of iterations and cost incurred to find the initial 'guess' maximum objective value of 30.5 for various initial training set sizes i.e. RVE results. The parametric study was carried out using the NCC decision-making policy. The final objective values obtained after reaching 300 iterations are also presented.

RVEs used to train		Iterations to exceed known max	Cost (ks)		Final objective value reached
Initial	Optimization		Initial	Optimization	
10	14	141	72	110.811	31.6071
20	17	171	144	137.106	31.3176
40	8	81	288	60.921	32.6166
70	8	81	504	60.921	32.7012
100	5	51	720	37.326	32.0356

Figure 3 compares the performance of three decision-making policies described in Section 2.4. To this end, the maximum objective value found by each optimization process at end of an iteration is plotted. As seen in Fig. 3, after about 50 iterations, all approaches come up with better design points resulting in higher objective values. Basically, after a

certain number of evaluations from lower fidelity models, the system is learning about the location of the maximum of the 'ground truth' model and evolves toward more and more optimal design points.

We note from this figure that the performance of both the NCC and CC-IC policies is similar. Out of all three policies, the NCC policy explores more since it is the least conservative. The CC-IC policy, on the other hand, prefers to exploit the current system knowledge rather than to explore highly uncertain regions. The CC-CC policy, which considers the calculation of the full modeling cost in addition to the information source cost, affects the decision making process in that it shows slower convergence to the target properties. This is likely due to the variable number of iterations between RVE calls, which influences how extensively the framework is exploring the design space. Additionally, by defining the cost-controlled calls to the RVE model, we let the system exhaust the cheap information sources before each call, allowing for more exploration of the design space. However, we balance the number of queries from the cheap information sources by considering the modeling cost as well in this policy. Note that for a fixed number of iterations, the CC-CC policy is not calling the RVE as much as the other policies. Thus, due to the fewer calls to the RVE, it makes sense that the CC-CC policy does not reach high objective values as fast as the NCC and CC-IC policies.

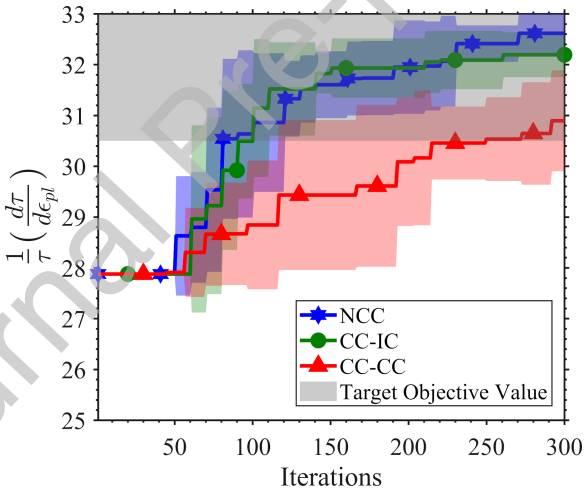


Figure 3: Comparing the maximum objective value found through the optimization process as a function of number of iterations for different decision-making policies: No Cost Consideration (NCC), Cost Consideration - Iteration Controlled (CC-IC) and Cost Consideration - Cost Controlled (CC-CC) 'ground truth' query.

Since the number of calls to the RVE and the number of information sources queried in each policy are different, every policy obtains an optimal solution (with respect to our target) at different overall cost. A comparison is done between all policies, Fig. 4, in order to show the objective value obtained as a function of the cost. The cost is calculated based on the time required to complete the defined number of iterations in kiloseconds (ks). As shown in Fig. 4, the CC-CC policy is suggesting *significantly better design points than* NCC and CC-IC policies *at the same cost*. As stated before, the advantage of the CC-CC policy

is that it favors the extraction of the most information out of a cheap source before querying the RVE. The CC-CC policy is thus the most cost-effective one, which is important in cases in which there are hard constraints the budget available to carry out the design. These cases are the norm rather than the exception in the context of materials design.

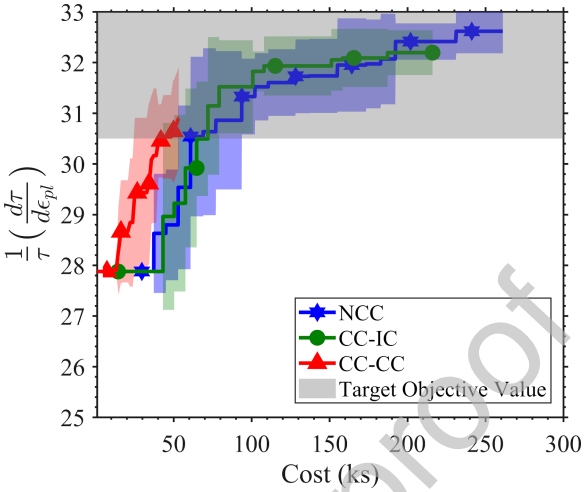


Figure 4: Comparing the maximum objective value found through the optimization process as a function of total cost for different decision-making policies: No Cost Consideration (NCC), Cost Consideration - Iteration Controlled (CC-IC) and Cost Consideration - Cost Controlled (CC-CC) 'ground truth' query.

To study the results in more detail, we investigate how each policy proceeds to an estimation of the optimal design point. Knowing that our framework employs multiple information sources to estimate the quantity of interest, the resource constraints on the decision-making process can have a large impact on how information sources are queried. By tracking the calls to different information sources during the optimization procedure, we can understand when and which information source had the highest value for Eq. (3) subject to the resource constraints of the given approach (CC-IC, CC-CC, and NCC). These results are shown in Fig. 5.

From Fig. 5a, it seems that the NCC policy favors the querying of the design space through the isostress reduced order model at the beginning of the optimization process. We note that this is not the cheapest source. Over time, the amount of knowledge gained from querying this source reaches a stage of diminishing returns. At this stage, the framework begins querying the isowork model—still not the cheapest source. As the optimization progresses, the NCC policy ends up encouraging the querying of the design space through all the available information sources. This is an expected result since there are no cost constraints and there is thus no penalty based on the information source queried. This behavior also reinforced the idea that this policy results in a more exploratory behavior, as the framework can utilize information from all sources more consistently. We note, however, that since the more expensive information sources are queried more compared to CC-IC and CC-CC policies, this approach will be considerably more expensive than the cost-constrained policies. A final observation to make is that, despite the fact that the micro-mechanical

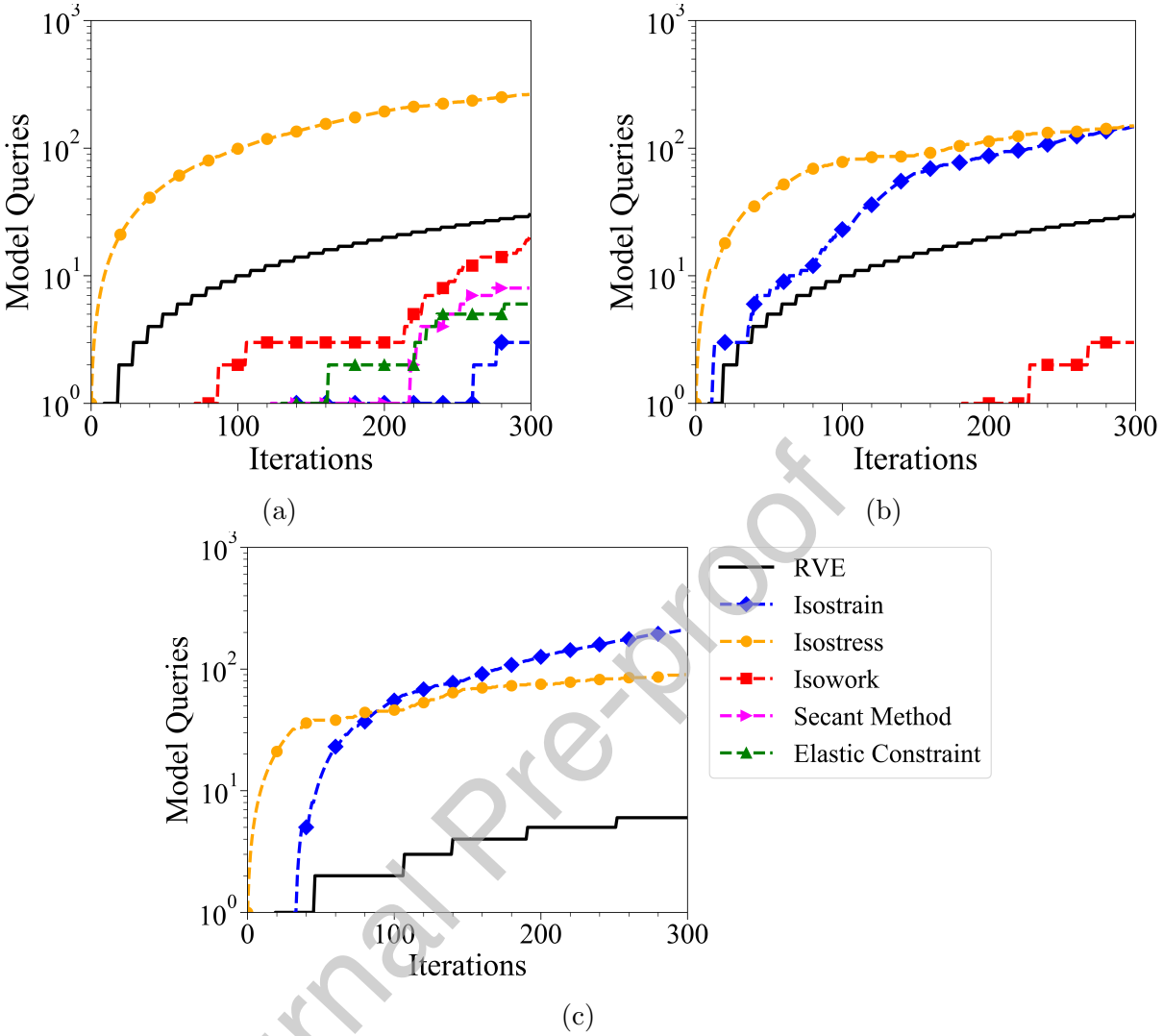


Figure 5: The progression of calls to different information sources as a function of the number of iterations for (a) No Cost Constraint (NCC), (b) Cost Constrained - Iteration Controlled (CC-IC) and (c) Cost Constrained - Cost Controlled (CC-CC) optimization procedures.

models (i.e. 'secant method' and 'elastic constraint') are more sophisticated than the simpler isostress approximation, the framework favors, by far, querying from this cheaper source, perhaps because the system has 'discovered' that it can make very useful inferences about the 'ground truth' by querying this cheap, and perhaps less accurate, source.

In a similar fashion to the NCC policy (Fig. 5a), the CC-IC (Fig. 5b) and CC-CC (Fig. 5c) policies favor querying the isostress information source first, followed by the isostrain approximation—these are the cheapest information sources available. Although the isostress approximation is more expensive than the isostrain model, the system initially finds more value in querying the former. After some iterations, however, the value—i,e, the ratio of information gained per cost—of the isostrain information source increases relative to

the isostress approximation. We note that when the framework follows the CC-CC policy (Fig. 5c), the transition in favor of the isostrain approximation (cheaper than the isostress model by a factor of ∼4) happens faster, as this is the most effective strategy to follow under hard resource constraints.

Considering the cost difference between the cheapest and most expensive information sources, the value of the information obtained from the more expensive information sources must be considerably higher to have a higher cost adjusted value in comparison to the cheaper information sources. As such, the framework is expected to exhaust all the information from the cheaper sources before moving on to the more expensive sources. In fact, it can be seen that none of the micro-mechanical models is queried at all under CC-IC and CC-CC and the isowork information source is only queried under CC-IC at the very last stages of the process.

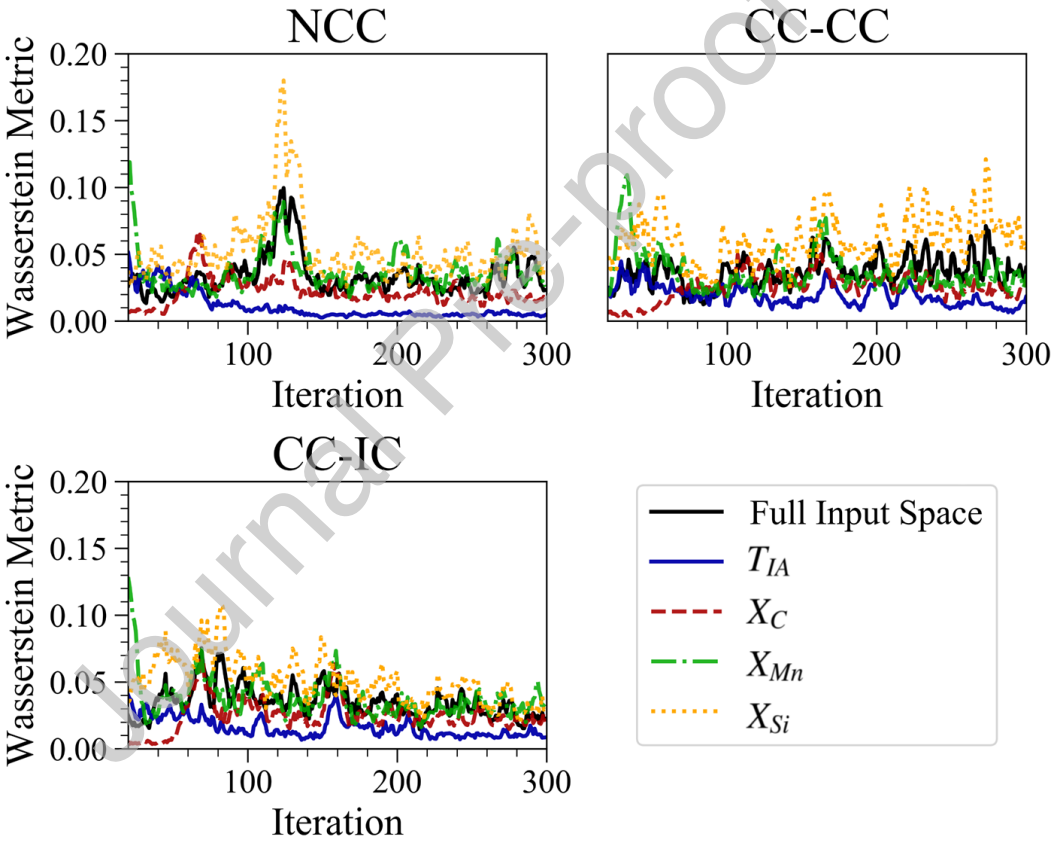


Figure 6: Evolution of the Wasserstein Metric for the full four-dimensional input space and each of the individual dimensions, temperature (T_{IA}), carbon (C), manganese (Mn) and silicon (Si) as a function of the number of iterations for No Cost Constraint (NCC), Cost Constrained - Cost Controlled (CC-CC) and Cost Constrained - Iteration Controlled (CC-IC) optimization procedures.

A major challenge of any optimization framework is to detect when convergence has been reached. In this very specific case, the problem is even more challenging, as the evaluation of the 'ground truth' has a degree of noise arising from the stochastic nature of the process.

The framework is thus *expected to converge into a region, rather than a single point*, in the design space. One possible approach to observe how the framework is able to approach the optimal point (or region) in the design space is to utilize the Wasserstein metric, also referred to as the Earth mover's distance. The Wasserstein metric can be understood as a distance function between two probability distribution functions in a given metric space. A more complete description of the theory and properties of the Wasserstein metric can be found in ref. [34]. This metric has been used in many applications, such as comparing the color distribution in images [35, 36], measuring the mixing and convergence of Markov Chains [37, 38] and as the loss function for the training of Generative Adversarial Neural Networks [39].

In this work we compute the Wasserstein metric between multiple one-dimensional distributions to understand how the framework approaches the optimal point or region in the multi-dimensional design space. To this end, we first compute the Euclidean distance between the current best point (that corresponds to the maximum of the fused GP) and a reference point. The reference point is taken to be $\chi_0 = [0.3, 0.3, 0.3, 0.3]$ in the unit hypercube design space. As the optimization progresses, the evolving distance between the current best point and the reference point form a distribution. We use the distribution of the distance in the four-dimensional space as well as the distributions of its projection along the individual dimensions to compute the Wasserstein metric. Specifically, starting from iteration 20, we take the 20 distance measurements prior to the iteration, split them into two sets of 10 each, and then compare the two sets using the Wasserstein metric. As the process converges, the Wasserstein metric is expected to approach zero, provided the sample size is large enough. In our case, having 10 samples per distribution means that the metric will be noisy even close to convergence. Here, we consider that a measure of convergence corresponds to a condition in which the Wasserstein metric maintains an almost constant value over multiple iterations.

As can be seen in Fig. 6, the Wasserstein metric computed from the distribution of the distance in the full four-dimensional input space show that the three different optimization policies have similar convergence (or lack thereof) patterns. However, the evolution of the Wasserstein metric computed from the distributions of the projections of the distance along the individual dimensions does provide a better understanding of the optimization process. These show quite clearly that all three methods converge rapidly in the temperature 'dimension'. In other words, the framework finds the optimal temperature very quickly. The next design input to reach convergence is the amount of carbon in the steel, followed by the manganese concentration. We note, however, that the silicon signal is very noisy and there is no indication of convergence. The slower convergence of manganese and non-convergence in silicon is somewhat expected since both elements do not play a significant role in controlling the normalized strain hardening rate of the material.

Finally, after all the evaluations are done and the system has built the final fused models, the fused models mean functions are searched to find the highest objective value they suggest. However, the predicted objective value is not important in itself, but rather, the corresponding design point is more valuable. The optimal design point suggested by all three approaches is compared with the predictions of a single PSP chain involving thermodynamic

calculations coupled with microstructure-based finite element calculations (RVE) in Fig. 7. In the figure, the contour plots show the variation of the maximum objective value where the maximum is taken over two (hidden) design variables, χ_3 and χ_4, for a given set of values of the remaining two (displayed) design variables, χ_1 and χ_2, i.e. $\max\left[f_{\chi_1,\chi_2}\left(\chi_3,\chi_4\right)\right]$, with χ_1 and χ_2, as obtained from the predictions of the aforementioned single PSP chain. Similarly, two out of four optimal design variables suggested by all three approaches are plotted at a time in Fig. 7. As expected, the CC-CC policy still makes predictions slightly different than the NCC and CC-IC policies. This may be because with a fewer number of queries from the RVE, the system has made less corrections to the fused model prediction compared to NCC and CC-IC policies. On the other hand, both the NCC and CC-IC policies are predicting the optimal design point very closely to the 'true' optimal design point. Furthermore, from the chemistry/processing (design) space exploration plots in Fig. 7 it can be noted that the true optimal solution is not unique. For example, in Fig. 7a, we see that an objective value of $(1/\tau)\left(\mathrm{d}\tau/\mathrm{d}\epsilon_{pl}\right) \approx 32$ has multiple solution in temperature and carbon space, i.e. the optimal solution is a region in the temperature and carbon space. Similarly, in Figs. 7b and 7c we see that the solution in temperature and manganese space, and temperature and silicon space, respectively, is also not unique.

4. Discussion

In this work, we expanded on our prior work [13, 14] on multi-information source fusion BO framework to account for the material design space that are truly amenable to modifications without concerns on whether they are feasible or not. Note that while promising, a microstructure discovered by only considering material microstructure as design space can be optimal in principle but may not be accessible in practice. We have demonstrated our framework by optimizing the performance of a material with material chemistry and processing conditions as the design space, thus truly realizing the ICME paradigm. Specifically, we have utilized our material design framework to optimize the strength normalized strain hardening rate (a single metric that provides an indication of the ductility and formability) of a dual-phase (ferrite-martensite) material by adjusting the composition and heat-treatment temperature of the material. The framework utilized a surrogate model of the thermodynamic results to predict the phase volume fraction and composition in the material microstructure post single-stage heat-treatment (intercritical annealing followed by quenching), which was then used to predict the mechanical properties of the dual-phase material using a variety of micromechanical models (as cheap information sources) and a high throughput microstructure-based finite element model (as expensive 'ground truth').

The material design framework developed herein was tested under three different conditions. These conditions were shown to affect how quickly the framework attained a sufficiently high value of the objective function and the information sources that were used to achieve that solution. For example, it was seen that when comparing the NCC (no cost considetation) to the CC-IC (cost consideration-iteration controlled) or CC-CC (cost consideration-cost controlled) policies, the NCC policy was able to utilize all the information sources while the other policies only favored queries from the cheaper models. The results

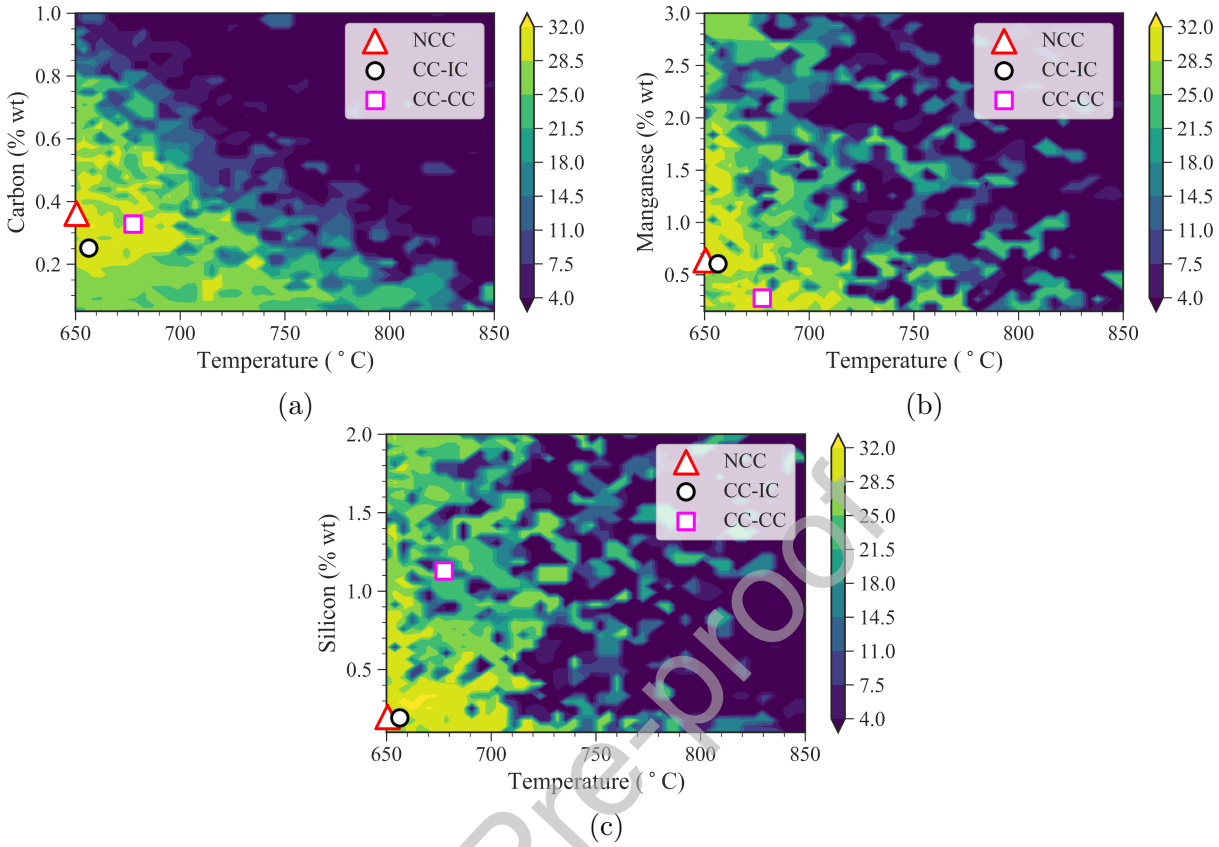


Figure 7: The optimal design point corresponding to the maximum objective value found by the three optimization processes, No Cost Constraint (NCC), Cost Constrained - Cost Controlled (CC-CC) and Cost Constrained - Iteration Controlled (CC-IC), are overlaid on the contour plots of the normalized strain hardening rate (objective) in the design space (temperature, carbon, manganese and silicon) generated by extensively exploring the 'ground truth' process-structure-property modeling chain involving thermodynamic calculations coupled with microstructure-based finite element calculations (RVE).

indicate that querying extensively from all models leads to an increased exploration of the design space. This, however, may not be desired for real life design problems where it is often necessary to obtain an optimum material composition within a tight budget constraint. In this regard, it is worth mentioning that when studying different policies by comparing the amount of cost they have incurred and the objective values they have obtained, the CC-CC policy showed significantly better performance by using less resources compared to the other two policies. This highlights the value in extensively extracting information out of cheaper information sources.

One big challenge in the optimization of an unknown design space is how to ascertain that the solution obtained from the optimization framework has converged. To address this challenge, we utilized the distance of the suggested optimal design point to a fixed reference point at every iteration in order to track the convergence of the framework. This convergence is strongly tied to the influence that a given dimension(s) in the input space has on the final objective function. For example, in problem at hand, temperature and

carbon have the biggest influence on the objective function and it is to be expected that the solution subspace corresponding to these to input variables converges relatively fast. On the other hand, silicon and manganese only have a weak influence on the objective, thus it is expected that a rather broad range of their values can lead to similar values of the objective function—the design space is almost degenerate with respect to silicon and manganese, once an optimal region within the temperature-carbon 'slice' has been identified. This makes the optimization of those two parameters extremely challenging, and highly 'oscillatory'.

The fact that not all the dimensions in the problem are equally influential (recall Fig. 6) points to a possible strategy to make this framework even faster and more efficient: *find and follow directions (i.e. subspaces) in the design space along which the objective function changes the most*. This approach, known as the active subspace method [40], decomposes the design space in such a way that at any time only the most 'influential' or 'active' subspaces are considered when trying to find the optimum in a multi-dimensional space. Effectively, this approach would reduce the dimension of the problem, accelerating the rate at which the solution is approached. As hinted above, in this problem we would expect that such an active subspace search would hone in on the temperature-carbon subspace prior to refining the optimization by exploiting the manganese and/or silicon coordinates. As of this writing, the authors of the present work are developing and testing a multi-information source fusion BO framework with active subspace search, to be presented in a forthcoming work.

5. Concluding remarks

We have presented a materials design framework for optimizing the mechanical properties of a dual-phase material using a multi-information source fusion approach. This framework utilized a thermodynamic model in order to predict the phase constitution (a stand-in for microstructure information) of a dual phase steel, and then used a variety of micro-mechanical models with varying degrees of fidelity and cost, to predict its mechanical behavior under tensile loading. The method was shown to be robust when subjected to various resource constraints that imposed restrictions on how the framework would query the expensive high-fidelity 'ground truth'. It was found that these constraints played an important role on how different information sources were queried by the framework. The results demonstrate that in the case where we do not consider the actual computational cost of each information source and do not impose any budgetary constraint, the framework has a tendency to exhaustively explore the available information sources. On the other hand, when we impose tight budget constraints, the framework extensively extracts the information out of cheaper information sources and is still able to provide a reasonable prediction of the optimum design space.

While the proposed materials design framework has been demonstrated to be extremely robust in realizing the ICME paradigm, this work instigates several possible future directions. For example, the current work employs a simple single information source along the chemistry/process - structure chain. An obvious next step will be to also incorporate more sophisticated thermodynamics/kinetics-based modeling schemes (perhaps implemented through phase-field simulations [41]) to not only predict the phase volume fraction

and composition in the material microstructure but also grain size and morphology of the constituent phases. Another possible future research direction will be to incorporate the active subspace method [40] in order to take advantage of the fact that in most materials design problems—and engineering design problems in general—the objective function is more sensitive to some input variables compared to others. Exploiting such information can lead to even faster and more efficient materials optimization schemes.

Acknowledgements

The authors acknowledge the financial support provided by the U.S. National Science Foundation through grant No.NSF-CMMI-1663130, *DEMS: Multi-Information Source Value of Information Based Design of Multiphase Structural Materials*. DA and RA also acknowledge the grant No. NSF-DGE-1545403. RA also acknowledges the grant nos. NSF-IIS-1849085 and NSF-IIS-1835690.

References

[1] J. Allison, Integrated computational materials engineering: A perspective on progress and future steps, JOM 63 (4) (2011) 15–18.

[2] W. Y. Wang, B. Tang, D. Lin, C. Zou, Y. Zhang, S.-L. Shang, Q. Guan, J. Gao, L. Fan, H. Kou, et al., A brief review of data-driven icme for intelligently discovering advanced structural metal materials: Insight into atomic and electronic building blocks, Journal of Materials Research 35 (8) (2020) 872–889.

[3] B. Gautham, R. Kumar, S. Bothra, G. Mohapatra, N. Kulkarni, K. Padmanabhan, More efficient icme through materials informatics and process modeling, in: Proceedings of the 1st World Congress on Integrated Computational Materials Engineering (ICME), Wiley Online Library, 2011, p. 35.

[4] M. F. Horstemeyer, Integrated Computational Materials Engineering (ICME) for metals: using multiscale modeling to invigorate engineering design with science, John Wiley & Sons, 2012.

[5] R. Arróyave, D. L. McDowell, Systems approaches to materials design: past, present, and future, Annual Review of Materials Research 49 (2019) 103–126.

[6] D. L. McDowell, S. R. Kalidindi, The materials innovation ecosystem: a key enabler for the materials genome initiative, Mrs Bulletin 41 (4) (2016) 326.

[7] P. Voorhees, G. Spanos, et al., Modeling across scales: a roadmapping study for connecting materials models and simulations across length and time scales, Tech. rep., Tech. rep., The Minerals, Metals & Materials Society (TMS) (2015).

[8] V. Savic, L. Hector, U. Basu, A. Basudhar, I. Gandikota, N. Stander, T. Park, F. Pourboghrat, K. S. Choi, X. Sun, et al., Integrated computational materials engineering (ICME) multi-scale model development for advanced high strength steels, Tech. rep., SAE Technical Paper (2017).

[9] L. Lin, W. Ren, An implementation of icme in materials information exchanging interfaces, Materials discovery 12 (2018) 9–19.

[10] A. Talapatra, S. Boluki, T. Duong, X. Qian, E. Dougherty, R. Arróyave, Autonomous efficient experiment design for materials discovery with bayesian model averaging, Physical Review Materials 2 (11) (2018) 113803.

[11] A. Solomou, G. Zhao, S. Boluki, J. K. Joy, X. Qian, I. Karaman, R. Arróyave, D. C. Lagoudas, Multiobjective bayesian materials discovery: Application on the discovery of precipitation strengthened niti shape memory alloys through micromechanical modeling, Materials & Design 160 (2018) 810–827.

[12] B. Peherstorfer, K. Willcox, M. Gunzburger, Survey of multifidelity methods in uncertainty propagation, inference, and optimization, Siam Review 60 (3) (2018) 550–591.

[13] S. F. Ghoreishi, A. Molkeri, A. Srivastava, R. Arroyave, D. Allaire, Multi-information source fusion and optimization to realize icme: Application to dual-phase materials, Journal of Mechanical Design 140 (11).
[14] S. F. Ghoreishi, A. Molkeri, R. Arróyave, D. Allaire, A. Srivastava, Efficient use of multiple information sources in material design, Acta Materialia 180 (2019) 260–271.
[15] D. Allaire, K. Willcox, Fusing information from multifidelity computer models of physical systems, in: Information Fusion (FUSION), 2012 15th International Conference on, IEEE, 2012, pp. 2458–2465.
[16] W. D. Thomison, D. L. Allaire, A model reification approach to fusing information from multifidelity information sources, in: 19th AIAA Non-Deterministic Approaches Conference, 2017, p. 1949.
[17] P. I. Frazier, W. B. Powell, S. Dayanik, A knowledge-gradient policy for sequential information collection, SIAM Journal on Control and Optimization 47 (5) (2008) 2410–2439.
[18] R. Couperthwaite, D. Allaire, R. Arryave, Utilizing gaussian processes to fit high dimension thermodynamic data that includes estimated variability, Computational Materials Science (2020) 110133.
[19] D. Koistinen, R. Marburger, A general equation prescribing the extent of the austenite-martensite transformation in pure iron-carbon alloys and plain carbon steels, Acta Metallurgica 7 (1) (1959) 59–60.
[20] A. Srivastava, A. Bower, L. Hector Jr, J. Carsley, L. Zhang, F. Abu-Farha, A multiscale approach to modeling formability of dual-phase steels, Modelling and Simulation in Materials Science and Engineering 24 (2) (2016) 025011.
[21] A. Srivastava, H. Ghassemi-Armaki, H. Sung, P. Chen, S. Kumar, A. F. Bower, Micromechanics of plastic deformation and phase transformation in a three-phase trip-assisted advanced high strength steel: Experiments and modeling, Journal of the Mechanics and Physics of Solids 78 (2015) 46–69.
[22] D. Gerbig, A. Srivastava, S. Osovski, L. G. Hector, A. Bower, Analysis and design of dual-phase steel microstructure for enhanced ductile fracture resistance, International Journal of Fracture (2017) 1–24.
[23] W. Voigt, On the relation between the elasticity constants of isotropic bodies, Ann. Phys. Chem 274 (1889) 573–587.
[24] A. Reuss, Berechnung der fließgrenze von mischkristallen auf grund der plastizitätsbedingung für einkristalle., ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik 9 (1) (1929) 49–58.
[25] O. Bouaziz, P. Buessler, Mechanical behaviour of multiphase materials: an intermediate mixture law without fitting parameter, Revue de Métallurgie–International Journal of Metallurgy 99 (1) (2002) 71–77.
[26] G. Weng, The overall elastoplastic stress-strain relations of dual-phase metals, Journal of the Mechanics and Physics of Solids 38 (3) (1990) 419–441.
[27] D. M. Saylor, J. Fridy, B. S. El-Dasher, K.-Y. Jung, A. D. Rollett, Statistically representative three-dimensional microstructures based on orthogonal observation sections, Metallurgical and Materials Transactions A 35 (7) (2004) 1969–1979.
[28] Dassault Systemes, Vélizy-Villacoublay, France, ABAQUS user's manual (2017).
[29] S. F. Ghoreishi, D. Allaire, Multi-information source constrained bayesian optimization, Structural and Multidisciplinary Optimization 59 (3) (2019) 977–991.
[30] C. E. Rasmussen, C. K. I. Williams, Gaussian Processes for Machine Learning (Adaptive Computation and Machine Learning), The MIT Press, 2005.
[31] R. L. Winkler, Combining probability distributions from dependent information sources, Management Science 27 (4) (1981) 479–488.
[32] W. B. Powell, I. O. Ryzhov, Optimal learning, Vol. 841, John Wiley & Sons, 2012.
[33] P. Frazier, W. Powell, S. Dayanik, The knowledge-gradient policy for correlated normal beliefs, INFORMS journal on Computing 21 (4) (2009) 599–613.
[34] C. Villani, The Wasserstein distances, in: C. Villani (Ed.), Optimal Transport: Old and New, Springer Berlin Heidelberg, Berlin, Heidelberg, 2009, pp. 93–111.
[35] E. Reinhard, M. Adhikhmin, B. Gooch, P. Shirley, Color transfer between images, IEEE Computer Graphics and Applications 21 (5) (2001) 34–41.

[36] Y. Rubner, C. Tomasi, L. J. Guibas, The Earth Mover's Distance as a Metric for Image Retrieval, International Journal of Computer Vision 40 (2) (2000) 99–121.
[37] M.-F. Chen, Trilogy of Couplings and General Formulas for Lower Bound of Spectral Gap, in: L. Accardi, C. C. Heyde (Eds.), Probability Towards 2000, Springer New York, New York, NY, 1998, pp. 123–136.
[38] Y. Ollivier, Ricci curvature of markov chains on metric spaces (2007). arXiv:math/0701886.
[39] M. Arjovsky, S. Chintala, L. Bottou, Wasserstein Generative Adversarial Networks, in: D. Precup, Y. W. Teh (Eds.), Proceedings of Machine Learning Research, Vol. 70, PMLR, Proceedings of Machine Learning Research, 2017, pp. 214–223.
[40] S. F. Ghoreishi, S. Friedman, D. L. Allaire, Adaptive dimensionality reduction for fast sequential optimization with gaussian processes, Journal of Mechanical Design 141 (7) (2019) 071404.
[41] V. Attari, A. Cruzado, R. Arroyave, Exploration of the microstructure space in tialzrn ultra-hard nanostructured coatings, Acta Materialia 174 (2019) 459–476.

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.