The multifaceted nature of uncertainty in structure-property linkage with crystal plasticity finite element model

Anh Tran∗
Sandia National Laboratories, Albuquerque NM 87123.

Pieterjan Robbe†
Sandia National Laboratories, Livermore, CA 94550.

Tim Wildey‡
Sandia National Laboratories, Albuquerque NM 87123.

David Montes de Oca Zapiain§
Sandia National Laboratories, Albuquerque NM 87123.

Hojun Lim¶
Sandia National Laboratories, Albuquerque NM 87123.

Uncertainty quantification (UQ) plays a critical role in verifying and validating forward integrated computational materials engineering (ICME) models. Among numerous ICME models, the crystal plasticity finite element method (CPFEM) is a powerful tool that enables one to assess microstructure-sensitive behaviors and thus, bridge material structure to performance. Nevertheless, given its nature of constitutive model form and the randomness of microstructures, CPFEM is exposed to both aleatory uncertainty (microstructural variability), as well as epistemic uncertainty (parametric and model-form error). Therefore, the observations are often corrupted by the microstructure-induced uncertainty, as well as the ICME approximation and numerical errors. In this work, we highlight several ongoing research topics in UQ, optimization, and machine learning applications for CPFEM to efficiently solve forward and inverse problems.

The first aspect of this work addresses the UQ of constitutive models for epistemic uncertainty, including both phenomenological and dislocation-density-based constitutive models, where the quantities of interest (QoIs) are related to the initial yield behaviors. We apply a stochastic collocation (SC) method to quantify the uncertainty of the three most commonly used constitutive models in CPFEM, namely phenomenological models (with and without twinning), and dislocation-density-based constitutive models, for three different types of crystal structures, namely face-centered cubic (fcc) copper (Cu), body-centered cubic (bcc) tungsten (W), and hexagonal close packing (hcp) magnesium (Mg).

The second aspect of this work addresses the aleatory and epistemic uncertainty with multiple mesh resolutions and multiple constitutive models by the multi-index Monte Carlo method, where the QoI is also related to homogenized materials properties. We present a unified approach that accounts for various fidelity parameters, such as mesh resolutions, integration time-steps, and constitutive models simultaneously. We illustrate how multilevel sampling methods, such as multilevel Monte Carlo (MLMC) and multi-index Monte Carlo (MIMC), can be applied to assess the impact of variations in the microstructure of polycrystalline materials on the predictions of macroscopic mechanical properties.

The third aspect of this work addresses the crystallographic texture study of a single void in a cube. Using a parametric reduced-order model (also known as parametric proper orthogonal decomposition) with a global orthonormal basis as a model reduction technique, we demonstrate that the localized dynamic stress and strain fields can be predicted as a spatiotemporal problem.

∗Technical Staff Member, Scientific Machine Learning.
†Postdoctoral Appointee, Plasma & Reacting Flow Science.
‡Technical Staff Member, Scientific Machine Learning.
§Technical Staff Member, Computational Materials & Data Science.
¶Technical Staff Member, Computational Materials & Data Science.

The fourth aspect of this work highlights the constitutive model calibration using an optimization under microstructure-induced uncertainty with Bayesian optimization. To account for natural variability or the aleatory uncertainty of microstructure, we average the loss function over an ensemble of microstructures and couple the Monte Carlo estimator with an asynchronous parallel Bayesian optimization to calibrate a phenomenological constitutive model. The framework is demonstrated for 304L stainless steel.

The fifth aspect of this work solves a stochastic inverse problem in the structure-property relationship. In this aspect, we seek to consistently learn a distribution of microstructure features, in the sense that the forward propagation of this microstructure feature distribution through CPFEM matches a target distribution of homogenized materials properties.

I. Introduction

Process-structure-property relationship is the hallmark of materials science across multiple length-scales and time-scales. Along with experimental materials science, numerous integrated computational materials engineering (ICME) models have been developed over the last several decades to accurately predict and reliably quantify uncertainty for the prediction. The computational ICME approach and the emerging physics-informed and physics-constrained machine learning (ML) in materials science have significantly accelerated the materials design process to tailor materials properties depending on the need [1]. For materials science, uncertainty quantification (UQ) is an essential part, since microstructures are inherently noisy and can only be captured by statistical microstructure descriptors. Optimization (Opt) also plays an important role, mainly in calibrating ICME models to establish a predictive science process. The computer predictions with quantified uncertainty have long been visioned [2, 3], and it still holds true for ICME in computational materials science.

For structure-property relationship, phase-field and crystal plasticity finite element model (CPFEM) are arguably the most successful computational tools to numerically investigate different materials phenomena. In this paper, we highlight several on-going research efforts in UQ, optimization, and ML applications for CPFEM. The numerical implementations are mainly done through DREAM.3D [4] and DAMASK [5, 6]. The first code coupling was demonstrated in [7].

The remaining of the paper is organized as follows. Section II quantifies uncertainty with respect to three different constitutive models for initial yield behaviors. Section III applied multi-level/multi-index Monte Carlo (MLMC/MIMC) approach to quantify uncertainty for homogenized materials properties. Section IV develops a parametric reduced-order model using proper-orthogonal decomposition method to emulate the full-field stress-strain response of a void model. Section V demonstrates the asynchronous parallel Bayesian optimization to calibrate a phenomenological constitutive model for stainless steel 304L. Section VI applies a data-consistent stochastic inverse UQ technique to infer a distribution of microstructure features such that the forward propagation through CPFEM matches a target distribution of homogenized materials properties. Section VII discusses and concludes the paper.

II. UQ of constitutive models in CPFEM

In this section, we highlight our recent effort [8] in quantifying *epistemic* uncertainty associated with initial yield behaviors, which are characterized by yield strength ε_Y and yield stress σ_Y at 0.2% offset. In this approach, we impose a uniform prior for all constitutive model parameters. Three constitutive models are considered in this study: phenomenological constitutive model *with* and *without* twinning, dislocation-density-based constitutive model. Interested readers are referred to the DAMASK paper [6] (cf. Section 6) for more modeling description regarding for constitutive models.

The generalized Wiener-Askey polynomial chaos expansion [9] represents the second-order random process $f(\theta)$ as

$$
\begin{aligned}
f(\theta) \quad = \quad & c_0 I_0 + \sum_{i_1=1}^{\infty} c_{i_1} I_1(\xi_{i_1}(\theta)) \\
& + \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} c_{i_1 i_2} I_2(\xi_{i_1}(\theta), \xi_{i_2}(\theta)) \\
& + \sum_{i_1=1}^{\infty} \sum_{i_2=1}^{\infty} \sum_{i_3=1}^{\infty} c_{i_1 i_2 i_3} I_3(\xi_{i_1}(\theta), \xi_{i_2}(\theta), \xi_{i_3}(\theta)) + \cdots,
\end{aligned}
\tag{1}
$$

where $I_n(\xi_{i_1}, \cdots, \xi_{i_n})$ denotes the Wiener-Askey polynomial chaos of order n in terms of the random vector $\boldsymbol{\xi} = (\xi_{i_1}, \xi_{i_2}, \ldots, \xi_{i_n})$, and c's are polynomial chaos expansion coefficients to be determined. Without loss of generality,

Equation 1 can be rewritten as

$$f(\theta) = \sum_{j=0}^{\infty} \widehat{f}_j \mathbf{\Phi}_j(\boldsymbol{\xi}(\boldsymbol{\theta})),$$

where there is a one-to-one correspondence between the function $I_n(\xi_{i_1}, \cdots, \xi_{i_n})$ and $\mathbf{\Phi}_j(\boldsymbol{\xi})$.

Table 1 describes the relationship between the types of Wiener-Askey polynomial chaos and their corresponding underlying random variables. For uniformly distributed variables $\boldsymbol{\xi}$ used in this paper, the Wiener-Askey scheme [9] requires Legendre polynomials as the polynomial basis $\{\mathbf{\Phi}_j\}$.

Table 1 Relationship between the types of Wiener-Askey polynomial chaos and their underlying random variables θ

random variable	probability density function	polynomial	support range
Gaussian	$\dfrac{1}{\sqrt{2\pi}} e^{-\frac{\theta^2}{2}}$	Hermite	$(-\infty, \infty)$
uniform	$\dfrac{1}{2}$	Legendre	$[-1, 1]$
beta	$\dfrac{(1-\theta)^\alpha(1+\theta)^\beta}{2^{\alpha+\beta+1}B(\alpha+1,\beta+1)}$	Jacobi	$[-1, 1]$
gamma	$\dfrac{\theta^\alpha e^{-\theta}}{\Gamma(\alpha+1)}$	Laguerre	$[0, \infty)$

To mitigate the curse of dimensionality, we employ stochastic collocation method [10–12] to evaluate numerical integration on Gaussian abscissas and compute the polynomial chaos expansion coefficients using Smolyak sparse grid [13–16].

Following [17], [18], and [19], we summarize the variance-based global sensitivity analysis based on Sobol' decomposition as follows. In the spirit of generalized polynomial chaos expansion (i.e. Equation 1 after finite truncation), the Sobol' decomposition of $f(\boldsymbol{\xi})$ into the summands of increasing dimensions as

$$
\begin{aligned}
f(\xi_1, \ldots, \xi_n) &= f_0 + \sum_{i=1}^{n} \sum_{\alpha \in \mathcal{I}_1} f_\alpha \mathbf{\Phi}(\xi_i) \\
&+ \sum_{1 \le i_1 < i_2 \le n} \sum_{\alpha \in \mathcal{I}_{i_1 i_2}} f_\alpha \mathbf{\Phi}(\xi_{i_1}, \xi_{i_2}) + \cdots \\
&+ \sum_{1 \le i_1 < \cdots < i_s \le n} \sum_{\alpha \in \mathcal{I}_{i_1, \ldots, i_s}} f_\alpha \mathbf{\Phi}(\xi_{i_1}, \ldots, \xi_{i_s}) \\
&+ \cdots + \sum_{\alpha \in \mathcal{I}_{1,2,\ldots,n}} f_\alpha \mathbf{\Phi}(\xi_1, \ldots, \xi_n).
\end{aligned}
$$

Given a model of the form $y = f(\xi_1, \xi_2, \ldots, \xi_n)$, with y as a scalar, a variance-based first order effect for a generic factor ξ_i can be written as $\mathbb{V}_{\xi_i}\left[\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[y|\xi_i]\right]$, where $\boldsymbol{\xi}_{\sim i}$ is the vector $\boldsymbol{\xi}$ without the i-th element, i.e. $\boldsymbol{\xi}_{\sim i} = (\xi_1, \ldots, \xi_{i-1}, \xi_{i+1}, \ldots, \xi_n)$. The main effect sensitivity index (first-order sensitivity coefficient) is written as

$$S_i = \frac{\mathbb{V}_{\xi_i}\left[\mathbb{E}_{\boldsymbol{\xi}_{\sim i}}[y|\xi_i]\right]}{\mathbb{V}[y]}.$$

It is relatively well-known that

$$\mathbb{E}\left[\mathbb{V}\left[y|\boldsymbol{\xi}_{\sim i}\right]\right] + \mathbb{V}\left[\mathbb{E}\left[y|\boldsymbol{\xi}_{\sim i}\right]\right] = \mathbb{V}[y],$$

and therefore, the total effect sensitivity index can be obtained as

$$T_i = \frac{\mathbb{E}\left[\mathbb{V}\left[y|\boldsymbol{\xi}_{\sim i}\right]\right]}{\mathbb{V}[y]} = 1 - \frac{\mathbb{V}\left[\mathbb{E}\left[y|\boldsymbol{\xi}_{\sim i}\right]\right]}{\mathbb{V}[y]}.$$

For interested readers, more details and implementation are available in Tang et al. [20] and Crestaux et al. [18], where most of the computations are based on Monte Carlo sampling $\boldsymbol{\xi}$. In the context of this section, we can understand $\boldsymbol{\xi}$ as the set of parameters for the underlying constitutive model, whether it is phenomenological or dislocation-density-based, and y as the quantity of interest from the CPFEM model.
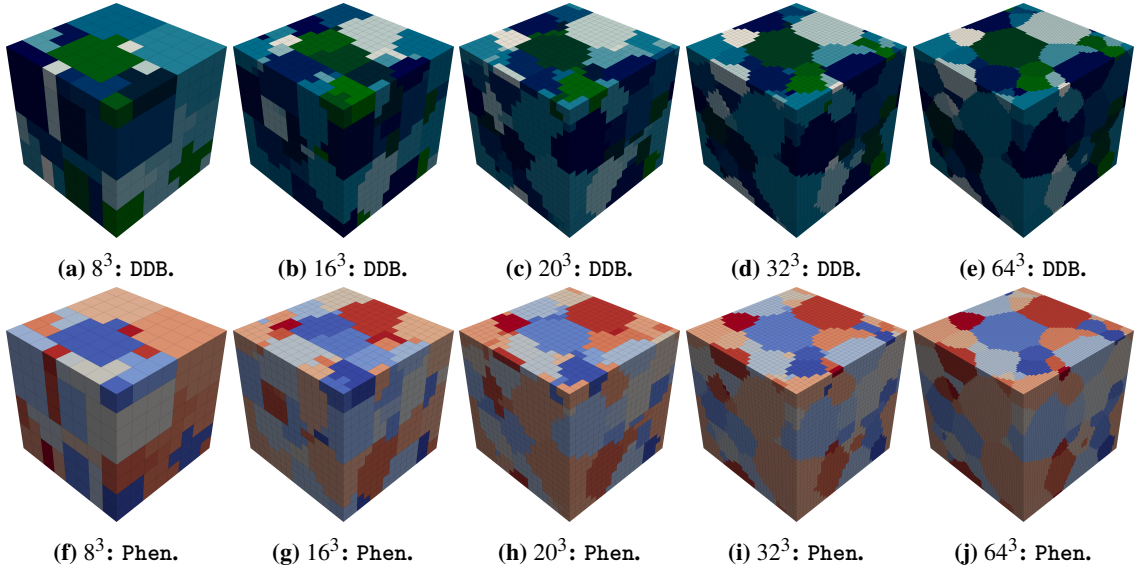
(a) fcc Cu.   (b) hcp Mg.   (c) bcc W.
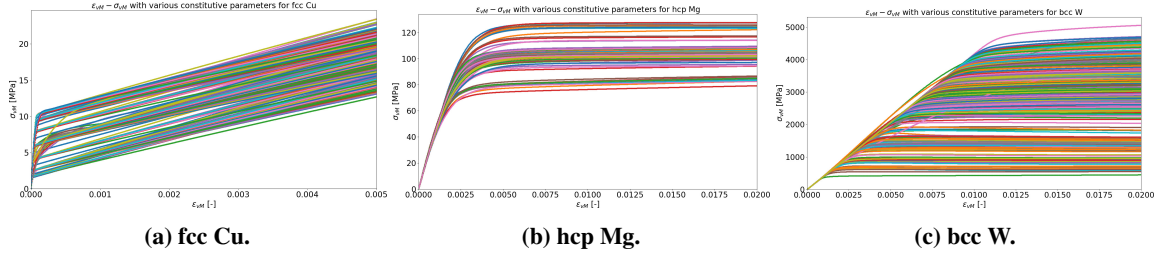
Fig. 1   Equivalent von Mises stress-strain plots. Reprinted from [8].

(a) Sobol' indices for σ_Y for fcc Cu.   (b) Sobol' indices for ε_Y for hcp Mg.   (c) Sobol' indices for ε_Y for bcc W.
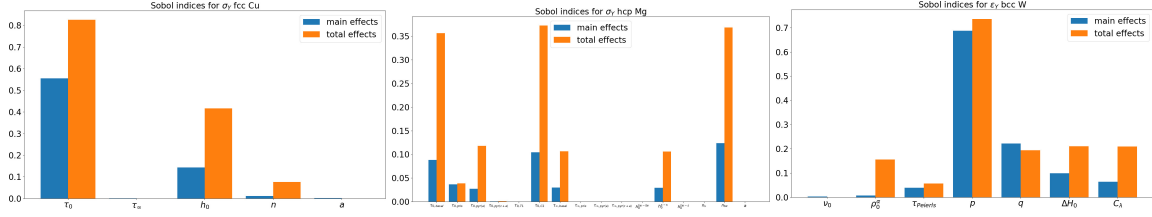
Fig. 2   Global sensitivity analysis with Sobol' indices for different constitutive models. Reprinted from [8].

## III. Multi-fidelity microstructure-induced UQ for CPFEM

The second aspect highlights our recent research effort in exploiting the well-posed fidelity hierarchy in CPFEM that is often overlooked in the literature to quantify *aleatory and epistemic* uncertainty. The *aleatory* uncertainty originates from the SERVE instantiation, whereas the *epistemic* uncertainty originates from the mesh discretization for SERVEs and the plasticity constitutive model. In this research thrust, we apply a relatively well-known Monte Carlo, called multi-level Monte Carlo [21, 22] and its multi-dimensional extension called multi-index Monte Carlo [23–25], to estimate a homogenized materials property in an efficient manner.

(a) 8^3: DDB.   (b) 16^3: DDB.   (c) 20^3: DDB.   (d) 32^3: DDB.   (e) 64^3: DDB.

(f) 8^3: Phen.   (g) 16^3: Phen.   (h) 20^3: Phen.   (i) 32^3: Phen.   (j) 64^3: Phen.

Fig. 3   A schematic illustration of multi-index Monte Carlo in CPFEM with 2-dimensional fidelity level. The first fidelity variable corresponds to the mesh resolution, whereas the second fidelity variable corresponds to the constitutive plasticity model. The first fidelity variable (i.e. mesh resolution) increases from the left to the right, where as the second fidelity variable (i.e. constitutive model) increases from the bottom to the top. The low-fidelity constitutive model is phenomenological constitutive model (`Phen`) and the high-fidelity constitutive model is dislocation-density-based (`DDB`) constitutive model.

4

In the MLMC case study, we consider an exemplar application of statistically equivalent representative volume elements (SERVEs), where the mesh discretization could be very coarse or very fine for α-Titanium. In the MIMC case study, we consider the extended version of the MLMC case study, where the first fidelity variable corresponds to the mesh resolution and the second variable corresponds to the constitutive models for an aluminum alloy. The phenomenological constitutive model is considered as the low-fidelity model, whereas the dislocation-density-based constitutive model is considered as the high-fidelity model. Figure 3 shows a schematic illustration of MIMC in two directions: the first direction – from left to right – corresponds to the mesh resolution, the second direction – from bottom to top – corresponds to the constitutive model.
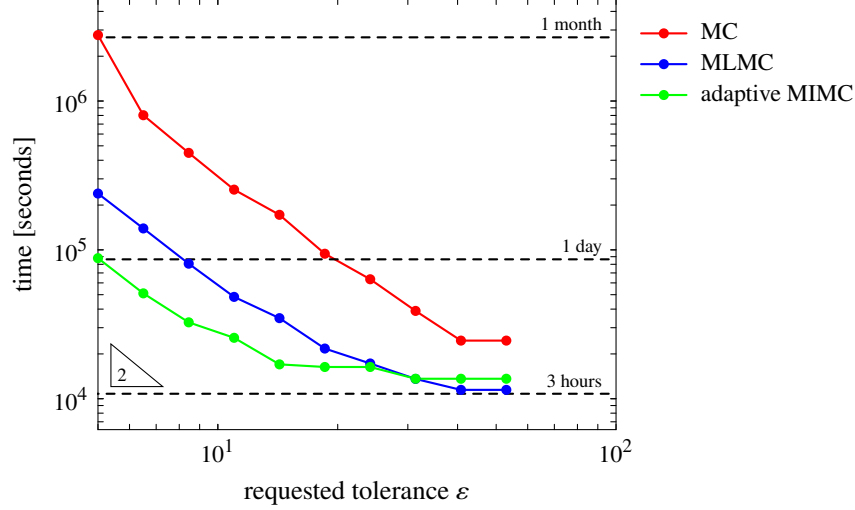
Fig. 4 **Cost of the adaptive MIMC and MLMC methods compared to the (estimated) complexity of the single-level MC method, expressed in terms of the total simulation time in seconds, as a function of the tolerance ε on the root mean-square error. For the target root mean-square error tolerance of $\varepsilon = 5$, the MLMC method is approximately 11.6 times faster than the MC method, and the adaptive MIMC method is 2.7 times faster than the MLMC method, resulting in a final speedup of adaptive MIMC over MC of $31.5\times$.**

Figure 4 compares the convergence behaviors of the Monte Carlo, MLMC, and adaptive MIMC estimators, respectively. For the MIMC case study, where the QoI is the effective Young modulus, with the root mean-square error $\varepsilon = 5$GPa, we demonstrate that the adaptive MIMC estimator is $2.7\times$ faster than the MLMC estimator, and the MLMC is $11.6\times$ faster than the MC estimator. By using an adaptive mesh resolution and an adaptive constitutive model, the homogenized materials properties can be estimated more efficiently and accurately with an unbiased estimator.

## IV. A parametric ROM for void model

In the third aspect, we develop a projection-based reduced-order model for single crystal void model using CPFEM, mainly following [26]. The idea behind is to decompose the random vector field $\mathbf{u}(\mathbf{x}, t; \mathbf{p})$, in the spirit of Karhunen-Loève expansion, into a set of deterministic spatial functions $\mathbf{l}_k(\mathbf{x})$ modulated by parameterized random time coefficients $\mathbf{q}(t; \mathbf{p})$ so that

$$\mathbf{u}(\mathbf{x}, t; \mathbf{p}) = \sum_{k=1}^{\infty} \mathbf{q}(t; \mathbf{p})\mathbf{l}_k(\mathbf{x}). \tag{2}$$

Originally introduced by Sirovich [27], the method of snapshots consider a set of snapshots $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m \in \mathbb{R}^n$ of state solutions computed at different instants in time and different orientations $\mathbf{p}$ (parametrized either as Euler angles (ρ_1, Φ, ρ_2) or generalized spherical harmonics coefficients, which is a robust and effective way to represent textures and orientations in polycrystalline materials [28, 29]), where $\mathbf{u}_j \in \mathbb{R}^n$ denotes the j-th snapshot and one collects $m < n$ snapshots, where n is typically large and in particular, (much) larger than the number of snapshots. Define the snapshot matrix $\mathbf{U} \in \mathbb{R}^{n \times m}$ whose jth column is the snapshot $\mathbf{u}_j$. The (thin) singular decomposition of $\mathbf{U}$ is written as

$$\mathbf{U} = \mathbf{L} \boldsymbol{\Sigma} \mathbf{R}^\top, \tag{3}$$

where $\mathbf{L} \in \mathbb{R}^{n \times m}$ and $\mathbf{R} \in \mathbb{R}^{m \times m}$ are the left and right singular vectors of $\mathbf{U}$, respectively, $\boldsymbol{\Sigma} = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_m) \in \mathbb{R}^{m \times n}$ is a rectangular orthogonal matrix (i.e. $\mathbf{L}^\top \mathbf{L} = \mathbf{R}^\top \mathbf{R} = \mathbf{I}_m$), where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_m$ are the singular values of
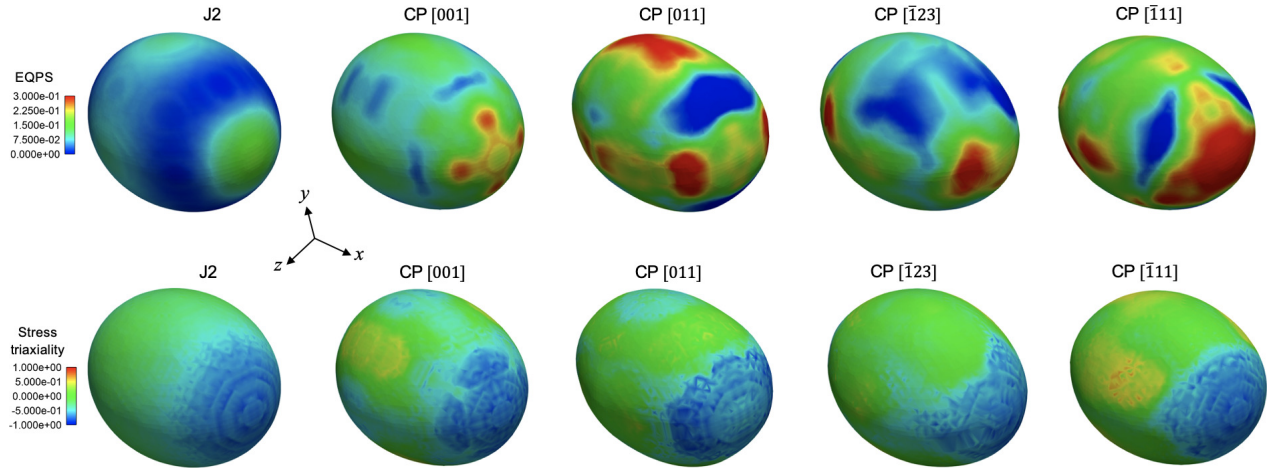
Fig. 5   Stress triaxiality for different crystal orientations.

(a) Reduced-order model.
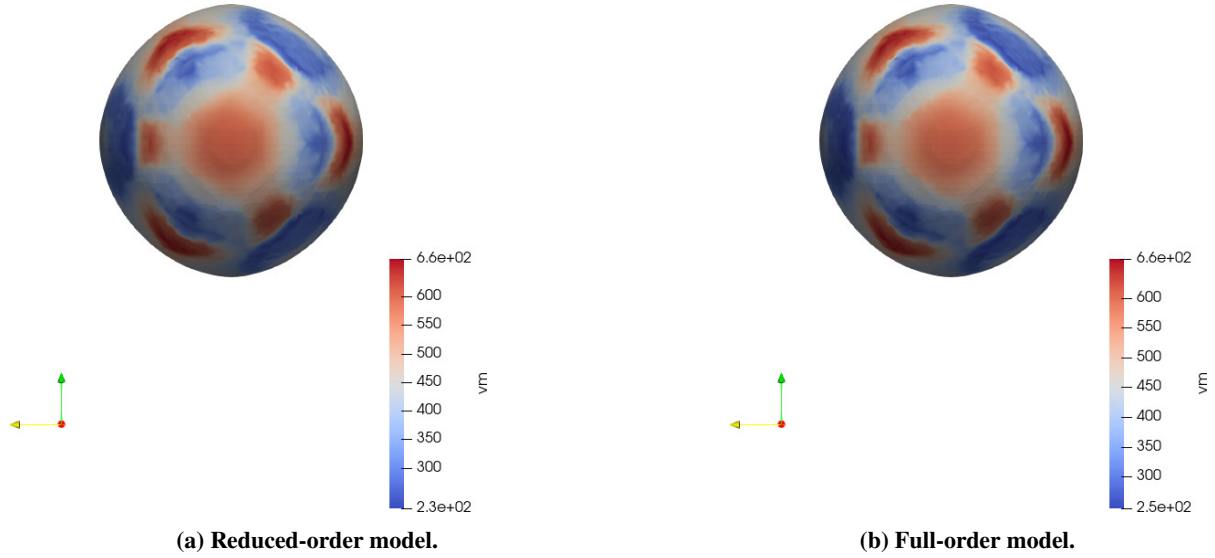
(b) Full-order model.

Fig. 6   Comparison between reduced-order model (Figure 6a) and full-order model (Figure 6b) in the $[\bar{1}11]$ direction shows an excellent agreement between reduced-order model and full-order model.

U. The proper orthogonal decomposition (POD) basis, $\mathbf{V}$, is chosen as the r left singular vectors of $\mathbf{L}$ that correspond to the r largest singular values. The POD basis is "optimal" in the sense that, for an orthonormal basis of size r, it minimizes the least squares error of snapshot reconstruction

$$\min_{\mathbf{V} \in \mathbb{R}^{n \times r}} \|\mathbf{U} - \mathbf{V}\mathbf{V}^\top \mathbf{U}\|_F^2 = \min_{\mathbf{V} \in \mathbb{R}^{n \times r}} \sum_{i=1}^{m} \|\mathbf{u}_i - \mathbf{V}\mathbf{V}^\top \mathbf{u}_i\|_F^2 = \sum_{i=r+1}^{m} \sigma_i^2$$

by the Eckart-Young theorem in Frobenius norm (cf. Theorem 2.4.8 and Section 2.5.2 [30]). It should be noted that the spatial modes of the direct POD are given by its left singular vectors $\mathbf{L}$, the temporal modes of the snapshot POD are given by its right singular vectors $\mathbf{R}$ since $\mathbf{U} \in \mathbb{R}^{n \times m}$. From Equation 3, let

$$\mathbf{Q} = \mathbf{\Sigma}\mathbf{R}^\top = \mathbf{L}^\top \mathbf{U},$$

where $\mathbf{Q} \in \mathbb{R}^{m \times m}$ then

$$\mathbf{U} = \mathbf{L}\mathbf{Q} = \sum_{k=1}^{m} l_k q_k,$$

where l_k is the kth column of $\mathbf{L}$ matrix, q_k is the kth row of $\mathbf{Q}$. We arrive at the discrete form of Equation 2 in temporal modes. Under the reduced-order model representation, the random vector field is approximated by a truncated Karhunen-Loève expansion as

$$\mathbf{u}(\mathbf{x}, t; \mathbf{p}) \approx \tilde{\mathbf{u}}(\mathbf{x}, t; \mathbf{p}) = \sum_{k=1}^{r} \mathbf{q}(t; \mathbf{p})\mathbf{l}_k(\mathbf{x}).$$

Now that the model can be represented by

$$\mathbf{U} = \mathbf{L}\mathbf{Q},$$

where $\mathbf{U} \in \mathbb{R}^{n \times m}, \mathbf{L} \in \mathbb{R}^{n \times m}, \mathbf{Q} \in \mathbb{R}^{m \times m}$ we can now interpolate the coefficient matrix $\mathbf{Q}$ to obtain

$$\mathbf{U}' = \mathbf{L}\mathbf{Q}',$$

where $\mathbf{U}' \in \mathbb{R}^{n \times m'}, \mathbf{L}' \in \mathbb{R}^{n \times m}, \mathbf{Q}' \in \mathbb{R}^{m \times m'}$. Following [27, 31], we construct a global basis approach to solve a parametric reduced-order model. One of the most common approaches in constructing the global basis matrices over the parameter space $\mathbf{p}$ is to concatenate the local basis matrices obtained by several parameter samples $\mathbf{p}_1, \ldots, \mathbf{p}_K$. Suppose that $\mathbf{Q}_1, \mathbf{Q}_2, \ldots, \mathbf{Q}_K$ denote the local basis matrices corresponding to $\mathbf{p}_1, \ldots, \mathbf{p}_K$, respectively, one can construct the global basis matrices $\mathbf{Q}$ using

$$\mathbf{Q} = [\mathbf{Q}_1, \mathbf{Q}_2, \ldots, \mathbf{Q}_K],$$

followed by an singular value decomposition (SVD) or a rank-revealing QR factorization to remove the rank-deficient components from $\mathbf{Q}$. Here we employ the Galerkin projection method [32, 33] to obtain the coefficient from the global basis in L_2-norm, exploiting the orthonormal properties of the global basis.

Figure 5 shows an exemplar of the stress triaxiality of a void using CPFEM for 9 different crystal orientations, along with a $J2$ plasticity model. For this anisotropic case study, the localized fields of the void (displacement in x, y, z directions, σ_{vM}, stress triaxiality) are a function of time and Euler angles (ρ_1, Φ, ρ_2). Figure 6 compares between a projection-based reduced-order model (ROM) and the full-order model (FOM, i.e. CPFEM) for the [$\bar{1}11$] orientation, which shows an excellent agreement between the ROM and the FOM.

## V. A high-throughput Bayesian optimization for constitutive model calibration

The fourth aspect addresses an effort in calibrating constitutive model in a high-throughput and asynchronous parallel manner, using one of our previous work [34–36]. In this section, we focus on the optimization under (microstructure-induced, also known as *aleatory*) uncertainty with applications to constitutive model calibration. The loss function to be minimized is measured in L_2-norm and normalized by the maximum observable equivalent strain ε_{vM}. A set of five constitutive parameters is used to parametrize a phenomenological constitutive model for stainless steel 304L. At any time, 12 CPFEM simulations are performed concurrently, where the batch configuration is set as (8,4,0). To account for the *aleatory* uncertainty, we average the loss function over an ensemble of 5 SERVEs, where the mesh of $5 \times 5 \times 5$ is used.

Figure 7 shows the comparison between the computational results produced by the optimal 5d constitutive parameters and the experimental data (marked as solid black line). The *aleatory* uncertainty is colored in a cyan shaded region (readers are referred to the color version online). Figure 7 shows a good agreement between experimental data and computational results up to approximately $\varepsilon \approx 0.42$ of strain, which is considerable for CPFEM. The optimal constitutive parameters is found after 352 iterations using the `aphBO-2GP-3B` algorithm [34].
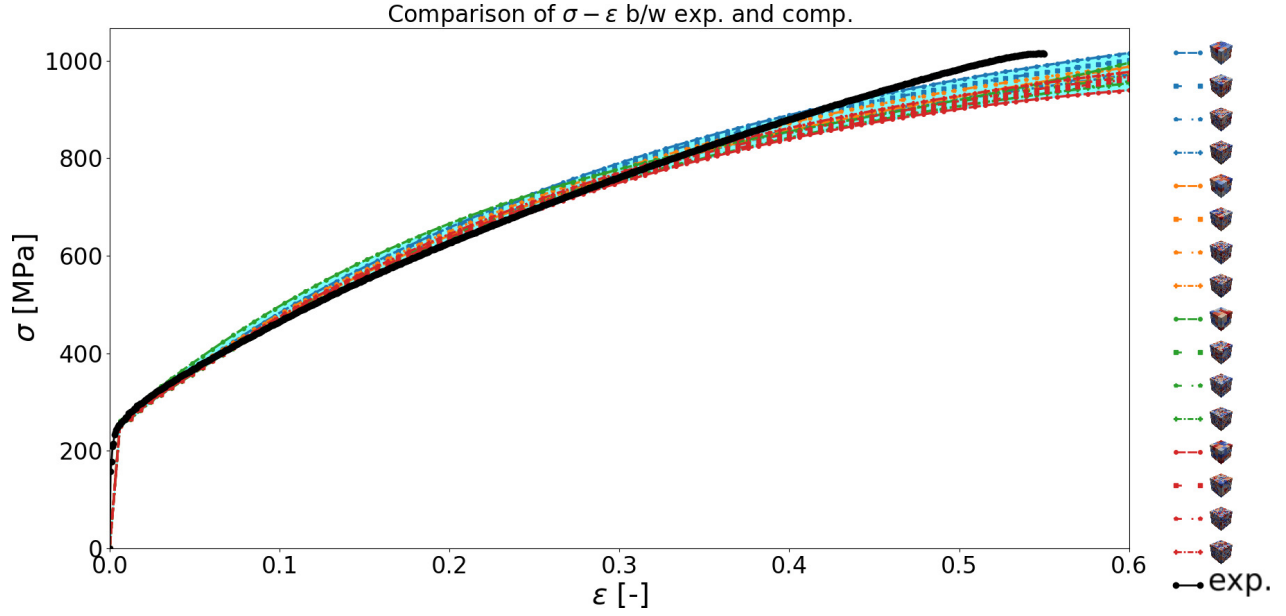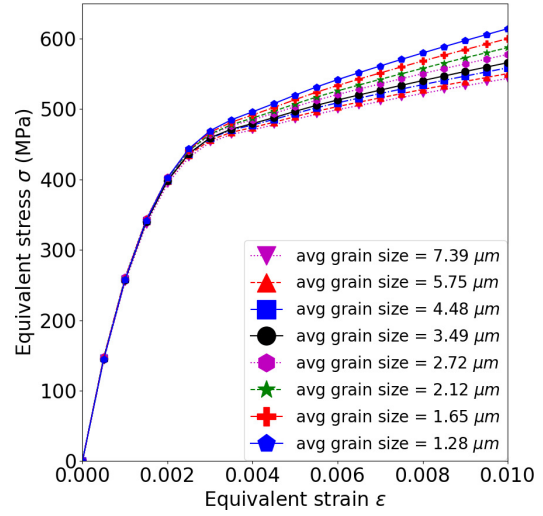
Fig. 7 Comparison of homogenized materials properties between experimental data and computational results for SS304L across different mesh resolutions and for different SERVEs.

VI. Solving stochastic inverse in property-structure relationship with ML
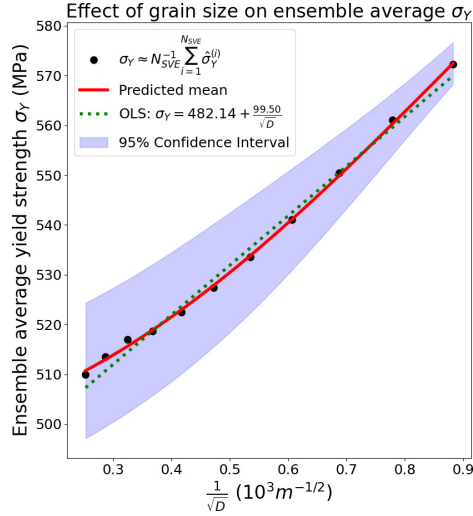
(a) Hall-Petch relationship by ordinary least square and Gaussian process regression.
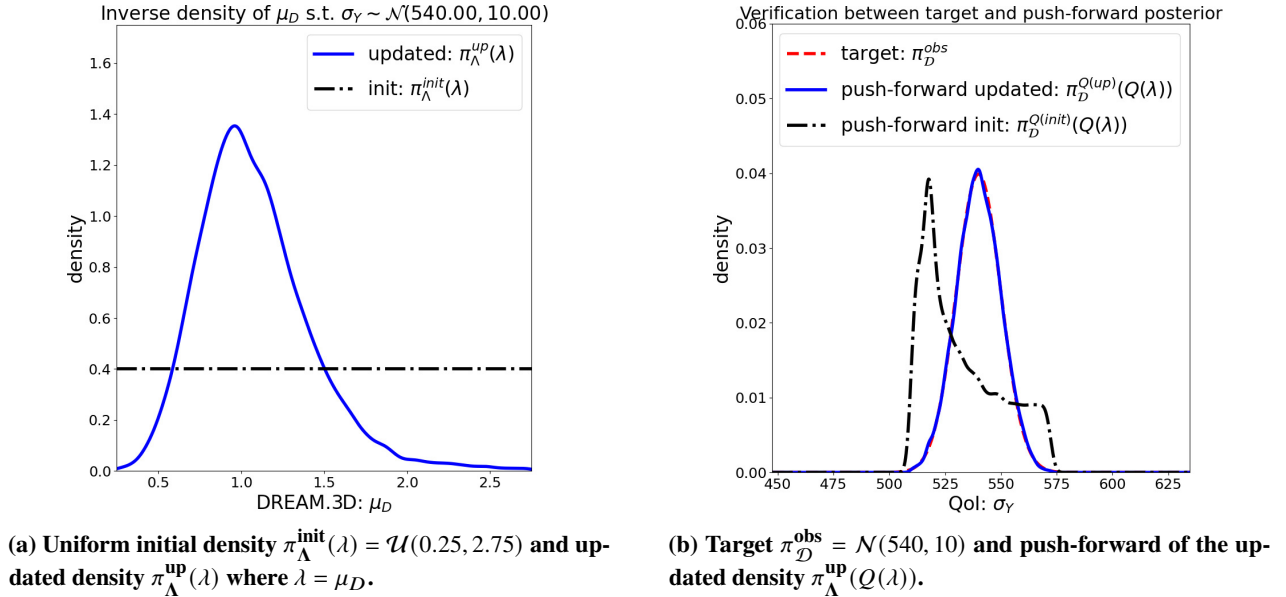

(b) Equivalent stress-strain curve for different average grain sizes.

Fig. 8 Figure 8a: The inferred ML Hall-Petch relationship with 95% confidence intervals for a TWIP steel, and a comparison with a least squares regression model. Figure 8b: The representative equivalent stress-strain relationship with respect to various average grain sizes. CPFEM simulations are performed with $64\mu m \times 64\mu m \times 64\mu m$ SVE at the loading condition of $\dot{\varepsilon}_{11} = 0.001\text{s}^{-1}$. (Reprinted with permission from [37] ©2020 The Minerals, Metals & Materials Society.)

The last aspect [37] applies a stochastic inverse UQ methodology [38–40] to infer a data-consistent distribution of microstructure features, in the sense that the push-forward distribution through CPFEM matches the target distribution of the homogenized materials properties of interests. In this case study, we employed a dislocation-density-based for twinning-induced plasticity/transformation-induced plasticity (TWIP/TRIP) Fe-22Mn-0.6C steels [41], where the density of average grain size is inferred based on a ML surrogate (i.e. a heteroscedastic Gaussian process regression)

for the Hall-Petch relationship to match target distribution of yield stress. The caveat in this case study is the notion of heteroscedastic behavior under a fixed size assumption of SERVE. Let D denotes the average grain size. When the average grain size D increases, the average grain volume scales as $O(D^3)$, and under the assumption of fixed volume for SERVEs, the number of grain size n decreases as $n \sim O(D^{-3})$. The variance of the Monte Carlo estimator scales as $n^{-1} \sim O(D^3)$. Hence larger average grain size would induce a larger variance in the Monte Carlo estimator. As such, the regression is heteroscedastic.

Figure 8a shows the Hall-Petch relationship obtained from CPFEM, estimated by the posterior mean of the heteroscedastic Gaussian process regression (——) and ordinary least square (·····). The ensemble average observations are denoted as •. Figure 8b shows the equivalent stress-strain curve for different average grain size at the initial yield 0.2% offset.

(a) Uniform initial density $\pi_\Lambda^{\mathbf{init}}(\lambda) = \mathcal{U}(0.25, 2.75)$ and updated density $\pi_\Lambda^{\mathbf{up}}(\lambda)$ where $\lambda = \mu_D$.

(b) Target $\pi_\mathcal{D}^{\mathbf{obs}} = \mathcal{N}(540, 10)$ and push-forward of the updated density $\pi_\Lambda^{\mathbf{up}}(Q(\lambda))$.

Fig. 9 **Figure 9a: The initial density, $\pi_\Lambda^{\mathbf{init}}(\lambda) = \mathcal{U}(0.25, 2.75)$, and updated density, $\pi_\Lambda^{\mathbf{up}}(\lambda)$, on the microstructure feature $\lambda = \mu_D$. Figure 9b: The target density on material properties, $\pi_\mathcal{D}^{\mathbf{obs}} = \mathcal{N}(540, 10)$, and push-forwards of the initial and updated densities. (Reprinted with permission from [37] ©2020 The Minerals, Metals & Materials Society.)**

Figure 9a shows the uniform *prior* and the *posterior* distributions of average grain size, where the *posterior* is updated using the stochastic inverse method. Figure 9b shows the normal target distribution $\mathcal{N}(540, 10)$, the push-forward of the *prior*, and the push-forward of the *posterior*. The push-forward of the *posterior* and the target distribution of the yield stress agrees very well with each other.

VII. Discussion & Conclusion

In this paper, we survey our recent or on-going research effort in UQ, optimization, and ML for structure-property relationship using CPFEM. As the microstructure is usually high-dimensional and naturally random, deploying UQ techniques is necessary to conduct an efficient and effective numerical studies. In this paper, we demonstrate multiple aspects of using UQ mathematical methodologies in the interest of CPFEM for structure-property relationship. Even though the field of UQ and multiscale computational materials science both have been found for a relatively long time, there is still plenty of research opportunities and open questions for further research. The objective of this paper is to highlight a few interesting and contemporary topics in UQ for CPFEM.

Acknowledgments

References

[1] de Pablo, J. J., Jackson, N. E., Webb, M. A., Chen, L.-Q., Moore, J. E., Morgan, D., Jacobs, R., Pollock, T., Schlom, D. G., Toberer, E. S., et al., "New frontiers for the materials genome initiative," *npj Computational Materials*, Vol. 5, No. 1, 2019, pp. 1–23.

[2] Oden, T., Moser, R., and Ghattas, O., "Computer predictions with quantified uncertainty, Part I," *SIAM News*, Vol. 43, No. 9, 2010, pp. 1–3.

[3] Oden, T., Moser, R., and Ghattas, O., "Computer predictions with quantified uncertainty, Part II," *SIAM News*, Vol. 43, No. 10, 2010, pp. 1–4.

[4] Groeber, M. A., and Jackson, M. A., "DREAM.3D: a digital representation environment for the analysis of microstructure in 3D," *Integrating materials and manufacturing innovation*, Vol. 3, No. 1, 2014, p. 5.

[5] Roters, F., Eisenlohr, P., Kords, C., Tjahjanto, D., Diehl, M., and Raabe, D., "DAMASK: the Düsseldorf Advanced MAterial Simulation Kit for studying crystal plasticity using an FE based or a spectral numerical solver," *Procedia Iutam*, Vol. 3, 2012, pp. 3–10.

[6] Roters, F., Diehl, M., Shanthraj, P., Eisenlohr, P., Reuber, C., Wong, S. L., Maiti, T., Ebrahimi, A., Hochrainer, T., Fabritius, H.-O., et al., "DAMASK–The Düsseldorf Advanced Material Simulation Kit for modeling multi-physics crystal plasticity, thermal, and damage phenomena from the single crystal up to the component scale," *Computational Materials Science*, Vol. 158, 2019, pp. 420–478.

[7] Diehl, M., Groeber, M., Haase, C., Molodov, D. A., Roters, F., and Raabe, D., "Identifying structure–property relationships through DREAM.3D representative volume elements and DAMASK crystal plasticity simulations: An integrated computational materials engineering approach," *JOM*, Vol. 69, No. 5, 2017, pp. 848–855.

[8] Tran, A., Wildey, T., and Lim, H., "Microstructure-sensitive uncertainty quantification for crystal plasticity finite element constitutive models using stochastic collocation method," *Frontiers in Materials*, Vol. 9, 2022, pp. 1–20.

[9] Xiu, D., and Karniadakis, G. E., "The Wiener–Askey polynomial chaos for stochastic differential equations," *SIAM Journal on Scientific Computing*, Vol. 24, No. 2, 2002, pp. 619–644.

[10] Babuška, I., Nobile, F., and Tempone, R., "A stochastic collocation method for elliptic partial differential equations with random input data," *SIAM Journal on Numerical Analysis*, Vol. 45, No. 3, 2007, pp. 1005–1034.

[11] Nobile, F., Tempone, R., and Webster, C. G., "A sparse grid stochastic collocation method for partial differential equations with random input data," *SIAM Journal on Numerical Analysis*, Vol. 46, No. 5, 2008, pp. 2309–2345.

[12] Xiu, D., "Fast numerical methods for stochastic computations: a review," *Communications in computational physics*, Vol. 5, No. 2-4, 2009, pp. 242–272.

[13] Novak, E., and Ritter, K., "High dimensional integration of smooth functions over cubes," *Numerische Mathematik*, Vol. 75, No. 1, 1996, pp. 79–97.

[14] Novak, E., and Ritter, K., "The curse of dimension and a universal method for numerical integration," *Multivariate approximation and splines*, Springer, 1997, pp. 177–187.

[15] Novak, E., and Ritter, K., "Simple cubature formulas with high polynomial exactness," *Constructive approximation*, Vol. 15, No. 4, 1999, pp. 499–522.

[16] Barthelmann, V., Novak, E., and Ritter, K., "High dimensional polynomial interpolation on sparse grids," *Advances in Computational Mathematics*, Vol. 12, No. 4, 2000, pp. 273–288.

[17] Sudret, B., "Global sensitivity analysis using polynomial chaos expansions," *Reliability engineering & system safety*, Vol. 93, No. 7, 2008, pp. 964–979.

[18] Crestaux, T., Le Maıtre, O., and Martinez, J.-M., "Polynomial chaos expansion for sensitivity analysis," *Reliability Engineering & System Safety*, Vol. 94, No. 7, 2009, pp. 1161–1172.

[19] Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., and Tarantola, S., "Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index," *Computer physics communications*, Vol. 181, No. 2, 2010, pp. 259–270.

[20] Tang, G., Iaccarino, G., and Eldred, M., "Global sensitivity analysis for stochastic collocation," *51st AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference 18th AIAA/ASME/AHS Adaptive Structures Conference 12th*, 2010, p. 2922.

[21] Giles, M. B., "Multilevel Monte Carlo path simulation," *Operations research*, Vol. 56, No. 3, 2008, pp. 607–617.

[22] Giles, M. B., "Multilevel Monte Carlo methods," *Acta Numerica*, Vol. 24, 2015, pp. 259–328.

[23] Haji-Ali, A.-L., Nobile, F., and Tempone, R., "Multi-index Monte Carlo: when sparsity meets sampling," *Numerische Mathematik*, Vol. 132, No. 4, 2016, pp. 767–806.

[24] Haji-Ali, A.-L., Nobile, F., Tamellini, L., and Tempone, R., "Multi-index stochastic collocation for random PDEs," *Computer Methods in Applied Mechanics and Engineering*, Vol. 306, 2016, pp. 95–122.

[25] Haji-Ali, A.-L., Nobile, F., Tamellini, L., and Tempone, R., "Multi-index stochastic collocation convergence rates for random PDEs with parametric regularity," *Foundations of Computational Mathematics*, Vol. 16, No. 6, 2016, pp. 1555–1605.

[26] Benner, P., Gugercin, S., and Willcox, K., "A survey of projection-based model reduction methods for parametric dynamical systems," *SIAM review*, Vol. 57, No. 4, 2015, pp. 483–531.

[27] Sirovich, L., "Turbulence and the dynamics of coherent structures. I. Coherent structures," *Quarterly of applied mathematics*, Vol. 45, No. 3, 1987, pp. 561–571.

[28] Bunge, H.-J., *Texture analysis in materials science: mathematical methods*, Elsevier, 2013.

[29] Kalidindi, S. R., *Hierarchical materials informatics: novel analytics for materials data*, Elsevier, 2015.

[30] Golub, G. H., and Van Loan, C. F., *Matrix computations*, Johns Hopkins University Press, 2013.

[31] Berkooz, G., Holmes, P., and Lumley, J. L., "The proper orthogonal decomposition in the analysis of turbulent flows," *Annual review of fluid mechanics*, Vol. 25, No. 1, 1993, pp. 539–575.

[32] Bui-Thanh, T., Willcox, K., and Ghattas, O., "Model reduction for large-scale systems with high-dimensional parametric input space," *SIAM Journal on Scientific Computing*, Vol. 30, No. 6, 2008, pp. 3270–3288.

[33] Carlberg, K., Barone, M., and Antil, H., "Galerkin v. least-squares Petrov–Galerkin projection in nonlinear model reduction," *Journal of Computational Physics*, Vol. 330, 2017, pp. 693–734.

[34] Tran, A., Eldred, M., Wildey, T., McCann, S., Sun, J., and Visintainer, R. J., "aphBO-2GP-3B: a budgeted asynchronous parallel multi-acquisition functions for constrained Bayesian optimization on high-performing computing architecture," *Structural and Multidisciplinary Optimization*, Vol. 65, No. 4, 2022, pp. 1–45.

[35] Tran, A., Sun, J., Furlan, J. M., Pagalthivarthi, K. V., Visintainer, R. J., and Wang, Y., "pBO-2GP-3B: A batch parallel known/unknown constrained Bayesian optimization with feasibility classification and its applications in computational fluid dynamics," *Computer Methods in Applied Mechanics and Engineering*, Vol. 347, 2019, pp. 827–852.

[36] Tran, A., Tranchida, J., Wildey, T., and Thompson, A. P., "Multi-fidelity machine-learning with uncertainty quantification and Bayesian optimization for materials design: Application to ternary random alloys," *The Journal of Chemical Physics*, Vol. 153, 2020, p. 074705.

[37] Tran, A., and Wildey, T., "Solving stochastic inverse problems for property-structure linkages using data-consistent inversion and machine learning," *JOM*, Vol. 73, 2020, pp. 72–89.

[38] Butler, T., Jakeman, J., and Wildey, T., "Convergence of Probability Densities Using Approximate Models for Forward and Inverse Problems in Uncertainty Quantification," *SIAM Journal on Scientific Computing*, Vol. 40, No. 5, 2018, pp. A3523–A3548.

[39] Butler, T., Jakeman, J., and Wildey, T., "Combining push-forward measures and Bayes' rule to construct consistent solutions to stochastic inverse problems," *SIAM Journal on Scientific Computing*, Vol. 40, No. 2, 2018, pp. A984–A1011.

[40] Butler, T., Wildey, T., and Yen, T. Y., "Data-consistent inversion for stochastic input-to-output maps," *Inverse Problems*, 2020.

[41] Wong, S. L., Madivala, M., Prahl, U., Roters, F., and Raabe, D., "A crystal plasticity model for twinning-and transformation-induced plasticity," *Acta Materialia*, Vol. 118, 2016, pp. 140–151.