SAND2020-3724J

Multiscale stochastic reduced-order model for uncertainty propagation using Fokker-Planck equation with microstructure evolution applications

Anh Tran[a,c], Jing Sun[b], Dehao Liu[a], Tim Wildey[c], Yan Wang[a,*]

[a] *Woodruff School of Mechanical Engineering, Georgia Institute of Technology, Atlanta, GA 30332*
[b] *Department of Population Health Science, Augusta University, Augusta, GA 30912*
[c] *Optimization and Uncertainty Quantification Department, Sandia National Laboratories, Albuquerque, NM 87123*

Abstract

Uncertainty involved in computational materials modeling needs to be quantified to enhance the credibility of predictions. Tracking the propagation of model-form and parameter uncertainty for each simulation step, however, is computationally expensive. In this paper, a multiscale stochastic reduced-order model (ROM) is proposed to propagate the uncertainty as a stochastic process with Gaussian noise. The quantity of interest (QoI) is modeled by a non-linear Langevin equation, where its associated probability density function is propagated using Fokker-Planck equation. The drift and diffusion coefficients of the Fokker-Planck equation are trained and tested from the time-series dataset obtained from direct numerical simulations. Considering microstructure descriptors in the microstructure evolution as QoIs, we demonstrate our proposed methodology in three integrated computational materials engineering (ICME) models: kinetic Monte Carlo, phase field, and molecular dynamics simulations. It is demonstrated that once calibrated correctly using the available time-series datasets from these ICME models, the proposed ROM is capable of propagating the microstructure descriptors dynamically, and the results agree well with the ICME models.

Keywords: stochastic reduced-order model, uncertainty propagation, Fokker-Planck equation, kinetic Monte Carlo, phase field, molecular dynamics, microstructure evolution

1. Introduction

Simulating the dynamic behaviors of material systems is one of the most important tasks for materials modeling. Quantities of interest (QoIs) are mostly related to the evolution of the systems along time. For instance, at atomistic scale, thermodynamic and mechanical properties of materials can be predicted by molecular dynamics (MD) simulation. At mesoscale, solidification processes are simulated by phase field (PF), kinetic Monte Carlo (kMC), and cellular automaton models. Given the model-form and parameter uncertainty associated with these models, the credibility of simulation predictions largely relies on how the uncertainty can be effectively quantified. Model-form uncertainty arises with simplification and approximation during model construction, whereas parameter uncertainty is associated with the parameter calibration process. The major challenges of uncertainty quantification (UQ) for materials modeling are associated with the high dimensionality of models and the complexity of uncertainty propagation during dynamics simulations. Here, the

*Corresponding author. Email: yan.wang@me.gatech.edu

complexity of uncertainty propagation problem is considered, where uncertainty associated with input parameters and initial conditions of material systems evolves along time during the dynamics simulation. The direct modeling of uncertainty propagation in material systems along directly simulated time steps suffers from the time-scale issue, because a very short time step is usually needed for each iteration in order to obtain the required fidelity of material systems and hundreds of thousands of iterations are typical. Regular UQ methods could add significant overhead costs to these simulations, which themselves already are computationally expensive to obtain meaningful QoIs.

Mathematical models have been developed to propagate uncertainty as stochastic processes. For example, the Kramers-Moyal expansion has been used to model the evolution of probability density functions (PDFs) of the QoI along time. When only the first and second orders of expansion are considered, the model is simplified to the Fokker-Planck equation. The Fokker-Planck equation is a deterministic ordinary differential equation of PDFs, which is equivalent to the stochastic differential equation usually employed to model Langevin dynamics. Generally speaking, many QoIs, but not all, in materials science can be modeled by a diffusion process, especially the QoIs that are related to diffusion problems. Such problems are common in the context of materials science, such as creep diffusion, boundary diffusion, dislocation diffusion, and so on. For example, grain growth is related to the diffusion of grain boundary, in a sense that the grain size stochastically increases as time advances. Another example is spinodal decomposition: similar to the grain growth, clusters corresponding to different phases in spinodal decomposition also rapidly grow and coalesce in the same manner with grains in the grain growth problem. Thus, it is natural to use diffusion processes to model microstructural evolution and its relevant QoIs, when it can be physically justified. One of the key assumptions in our work is that the QoI can be modeled by an one-dimensional diffusion process whose drift and diffusion coefficients can be estimated using the available time-series dataset, i.e. the temporal experimental or computational data. Such assumption is widely supported from various fields of science, where time-series data is involved, for example, neuroscience, cardiology, finance, economy, surface science, turbulence [1], seismic time-series and epileptic brain dynamics [2].

To alleviate the time-scale issue in propagating uncertainty in materials modeling, in this paper, a multiscale stochastic reduced-order model (ROM) is introduced. The evolution of PDFs associated with uncertain QoIs is modeled with the Fokker-Planck equation. The advantage of the proposed stochastic ROM approach is that the uncertainty propagation in the simulated material systems can be significantly accelerated because the time scale used in the Fokker-Planck equation is independent from the ones in the direct numerical simulations. The parameters of the stochastic ROM are the drift and diffusion coefficients. They need to be calibrated based on the direct numerical simulations. With well-calibrated parameters, the stochastic ROM can predict the evolution of QoIs. In this work, the distributions of QoIs from the direct numerical simulations as time series are divided into two time periods or stages. The data from the first stage are used to train and calibrate the ROM, whereas the second stage is used to test the ROM performance. Several further assumptions are made in using the Langevin equation to describe the QoI (cf. Section 4.2), which is considered as a stochastic process in this paper.

In the rest of this paper, we denote $\xi(t)$ as a one-dimensional QoI in an arbitrary ICME model,

where $\xi(t)$ is considered as a stochastic process and modeled using the Langevin equation. The outline of the paper is as follows. Section 2 reviews the literature on stochastic reduced-order model methods, time-upscaling via time-parallelization methods and applications, and scale-bridging and multiscale methods for computational materials science. Section 3 provides the background and derivation of the Fokker-Planck equation. The mathematical foundation of the proposed methodology and the numerical procedure in applying the proposed method is described in Section 4. Section 5 provides three examples of kMC, PF, and MD simulations. The advantages and limitations of the proposed methods are discussed in Section 6. Section 7 concludes the paper.

2. Literature reviews

We briefly review related works on stochastic ROM for solving uncertainty propagation problems (Section 2.1), parallelization in time-domain (Section 2.2), which is used as a time upscaling approach for solving partial differential equation, and scale-bridging methodology and applications (Section 2.3).

2.1. Stochastic reduced-order models

Some researchers have applied the Fokker-Planck equation as the stochastic ROM for system dynamics simulation. For instance, Grigoriu [3] constructed a stochastic ROM with simple random functions to approximate an arbitrary random functions, where statistical discrepancy are minimized. Sarkar et al. [4] applied the method of Grigoriu [3] to quantify uncertainty in a corroding system and compare against sampling-based approaches. Mignolet and Soize [5] proposed another stochastic ROM for both model and parameter uncertainty in a stochastic finite element approach. This nonparametric approach accounts for both model and parameter uncertainty, compared to only parameter uncertainty in Ghanem and Spanos [6]. Goudon and Monasse [7] modified Lifshitz-Slyozov equation based on Fokker-Planck equation and demonstrated the approach with polymer systems. Ganapathysubramanian and Zabaras [8] proposed a data-driven ROM and solved it through stochastic collation and variational multiscale methods for thermal diffusion in random heterogeneous media applications. Bhattacharjee and Matouš [9] proposed a ROM based on Isomap, a non-linear manifold learning technique, in concert with neural networks, for heterogeneous hyperelastic materials. Latypov and Kalidindi [10] proposed a ROM based on two-point statistics correlation for two-phase composite materials.

2.2. Parallelization for systems dynamics simulation

To improve the computational efficiency of system dynamics simulation, parallelism methods for time dimension have been introduced [11], including multiple shootings, domain decomposition and waveform relaxation, multigrid, and direct time parallel methods. Nievergelt [12] proposed the first shooting type approach by pure time decomposition. Chartier and Phillipe [13], Saha et al. [14], Mayday and Turicini [15], and Guillaume [16] further developed and analyzed the method, often referred to as parareal algorithm in the literature. The iterative domain decomposition in space-time domains also have received attentions from many researchers, including Lelarasmee [17], Gander et al. [18, 19]. Multigrid method is another iterative approach that is not naturally parallel, but

their components can be parallelized in the entire space-time domain [12]. Notable work includes Hackbusch [20], Lubich and Ostermann [21], Horton and Vandewalle [22], Emmett and Minion [23], Neumüller [24]. Direct solvers are the last in the four classes of time parallelism methods. Prior work incudes Miranker and Liniger [25], Axelson and Verwer [26], Womble [27], Worley [28], Sheen et al. [29], Maday and Ronquist [30], Christlieb et al. [31], Güttel [32]. Despite their success, the methods of parallelization in time described above only scale if the computational resources are sufficient. Consequently, time acceleration is achieved only if the high-performance computing infrastructure is available. As far as the authors' knowledge, there is no directly related multiscale methods with focus on time-upscaling problems for computational materials applications.

2.3. Scale-bridging and multiscale methods

Multiscale methods aiming at solving multi-physics problem multiple length-scale are available. For example, Yotov et al. [33, 34] proposed and implemented the multiscale mortar finite element method for second order elliptic equations. E and Engquist and collaborators [35, 36] proposed the heterogeneous multiscale methods to efficiently approximate the macroscopic state of the system when the microscopic model is readily available. Variational multiscale methods is another approach that has received much attention in solving computational fluid dynamics problems, for instance, Hughes et al [37, 38, 39]. Another approach that couples atomistic to continuum level is the Atomistic-to-Continuum methods [40, 41, 42], which has found many users in computational materials science. However, all of the aforementioned multiscale methods are rather limited to multiple length-scale, leaving time-scale issues unresolved.

3. Background on Fokker-Planck equation

In this section, the background of uncertainty propagation using the Kramers-Moyal expansion and Fokker-Planck equation is provided, as it is the backbone of our proposed methodology. The one-dimensional (1D) non-linear Langevin equation for stochastic variable $\xi(t)$, which is the QoI in this paper, is mathematically equivalent to the Fokker-Planck equation. Following the notation of Risken [43] and Frank [44], the 1D non-linear Langevin equation reads

$$\dot{\xi} = h(\xi, t) + g(\xi, t)\Gamma(t), \tag{1}$$

where the Langevin force $\Gamma(t)$ is assumed to be a Gaussian distributed, white noise term with zero mean and δ correlation function [45], i.e.

$$\langle \Gamma(t) \rangle = 0, \quad \langle \Gamma(t)\Gamma(t') \rangle = \delta(t - t'). \tag{2}$$

$\delta(t - t')$ is the Dirac-delta function, and the intensity of noise is 1 by convention [46]. If $h(\xi, t) = 0$ and $g(\xi, t) = 1$, Equation 1 describes the Wiener process. There are two alternative ways to interpret the drift and diffusion coefficients of the Fokker-Planck equation. The first is based on Itô calculus and the second is based on Stratonovich calculus, depending on the existence of spurious or noise-induced drift [43]. For Itô calculus, $D^{(1)}(x, t) = h(x, t)$, whereas for Stratonovich calculus, $D^{(1)}(x, t) = h(x, t) + \dfrac{\partial g(x, t)}{\partial x} D^{(2)}(x, t)$. For both Itô and Stratonovich calculus, $D^{(2)}(x, t) = g^2(x, t)$.

With Itô calculus, Equation 1 can be rewritten as $d\xi = h(\xi,t)dt + g(\xi,t)dW_t$, where $\xi = \xi(t)$, and W_t is the Wiener process, which is the source of randomness [46]. $\xi(t)$ is the QoI dependent on time, also referred to as the state variable. We restrict $\xi(t)$ to only one dimension, even though it can be generalized for higher dimensional QoIs. In the scope of this paper, Stratonovich calculus is used to interpret the stochastic process in Equation 1. The stochastic process described in Equation 1 is a Markov process with δ-correlated force; this Markov properties is destroyed if $\Gamma(t)$ is no longer δ-correlated [43].

The ordinary differential equation of PDFs based on the Kramers-Moyal expansion is a general model of the evolution of probability distributions. The Fokker-Planck equation is a special case when only the first- and second-order spatial derivatives are considered. The Kramers-Moyal expansion for the PDF $f(x,t)$ associated with stochastic variable $\xi(t)$ can be written as [43]

$$\frac{\partial f(x,t)}{\partial t} = \sum_{j=1}^{n} \left(\frac{-\partial}{\partial x}\right)^j D^{(j)}(x,t) f(x,t). \tag{3}$$

where n is the number of truncated terms in Kramers-Moyal expansion, $D^{(j)}(x,t)$ is the Kramers-Moyal expansion coefficient. The Pawula's theorem [43] states that the Kramers-Moyal expansion may stop either after the first term or after the second term; if it does not stop after the second term, then it must contain an infinite number of terms. With only the first two terms, the Kramers-Moyal expansion is reduced to the Fokker-Planck equation as

$$\frac{\partial f(x,t)}{\partial t} = -\frac{\partial}{\partial x}\left[D^{(1)}(x,t) f(x,t)\right] + \frac{\partial^2}{\partial x^2}\left[D^{(2)}(x,t) f(x,t)\right], \tag{4}$$

where $D^{(1)}$ and $D^{(2)}$ are the drift and diffusion coefficients, respectively, and both are spatio-temporal functions.

Assume the evolution of the stochastic variable $\xi(t)$, which is the QoI, can be modeled by the Fokker-Planck equation. This assumption holds if $\xi(t)$ obeys a Langevin equation with Gaussian δ-correlated noise, it can be shown that all coefficients other than drift and diffusion coefficients vanish, and the Kramer-Moyal expansion simply reduces to the Fokker-Planck equation (cf. [43], Section 1.2.7). The Fokker-Planck equation is sometimes referred to as the backwards or second Kolmogorov equation in literature [1]. Furthermore, for the case of one stochastic variable QoI $\xi(t)$ in Equation 1, it is always possible to convert the Langevin equation from a multiplicative noise force g to an additive noise force, through a transformation of variable (cf. [43], Section 3.3).

Theorem 1 ([45, 47, 48]). *Denote the n^{th} central moment as $M^{(n)}(x,t)$, then*

$$\frac{\partial}{\partial t} M^{(n)}(x,t) = n! D^{(n)}(x,t) + \mathcal{O}(\tau^2) \tag{5}$$

Thus, the Kramers-Moyal coefficients can be estimated from sampling the time-series data as

$$D^{(n)}(x,t) = \lim_{\tau \to 0} \frac{1}{n!\tau} \langle [\xi(t+\tau) - \xi(t)]^n \rangle \Big|_{\xi(t)=x}. \tag{6}$$

It is noted that Theorem 1 follows from the Taylor series expansion in deriving Kramers-Moyal expansion. The evolution of mean and variance can be modeled through the two following corollaries.

Corollary 1. *The expectation of* $\xi(t)$*, denoted as* $\mathbb{E}[\xi(t)]$*,*

$$\mathbb{E}[\xi(t)] := \int_{-\infty}^{\infty} x f(x,t) dx, \tag{7}$$

must satisfy

$$\frac{\partial}{\partial t}\mathbb{E}[\xi(t)] = \int_{-\infty}^{\infty} D^{(1)}(x,t) f(x,t) dx. \tag{8}$$

If the drift coefficient is a temporal function, only i.e. $D^{(1)}(x,t) = D^{(1)}(t)$*, then*

$$\frac{\partial}{\partial t}\mathbb{E}[\xi(t)] = D^{(1)}(t). \tag{9}$$

Corollary 2. *Assume that the drift coefficient is a temporal function, i.e.* $D^{(1)}(x,t) = D^{(1)}(t)$*, then the variance of* $\xi(t)$*, denoted as* $\mathrm{Var}[\xi(t)]$*,*

$$\mathrm{Var}[\xi(t)] := \int_{-\infty}^{\infty} (x - \mathbb{E}[\xi(t)])^2 f(x,t) dx, \tag{10}$$

must satisfy

$$\frac{\partial}{\partial t}\mathrm{Var}[\xi(t)] = 2\int_{-\infty}^{\infty} D^{(2)}(x,t) f(x,t) dx \tag{11}$$

If the diffusion coefficient is also a temporal function, i.e. $D^{(2)}(x,t) = D^{(2)}(t)$*, then*

$$\frac{\partial}{\partial t}\mathrm{Var}[\xi(t)] = 2D^{(2)}(t). \tag{12}$$

4. Methodology

In this section, the proposed stochastic ROM method with time upscaling is introduced. The organization of this section is as follows. Section 4.1 presents the problem statement in terms of mathematics, along with the assumptions in Section 4.2. In Section 4.3, we introduce the calibration and training procedure for the proposed stochastic ROM. Section 4.4 describes a numerical treatment via Tikhonov regularization to prevent divergent solutions. In Section 4.5, the finite difference method in solving the Fokker-Planck equation is discussed.

4.1. Problem statement

Figure 1 provides a schematic overview of the proposed stochastic ROM method. First, the ROM needs to be trained, where the drift and diffusion coefficients are calibrated using the predicted QoIs from direct numerical simulations of materials. After test and validation with additional materials simulation data, the stochastic ROM then can predict the uncertainty propagation efficiently with a much longer time step than the time steps used in the direct numerical simulations.

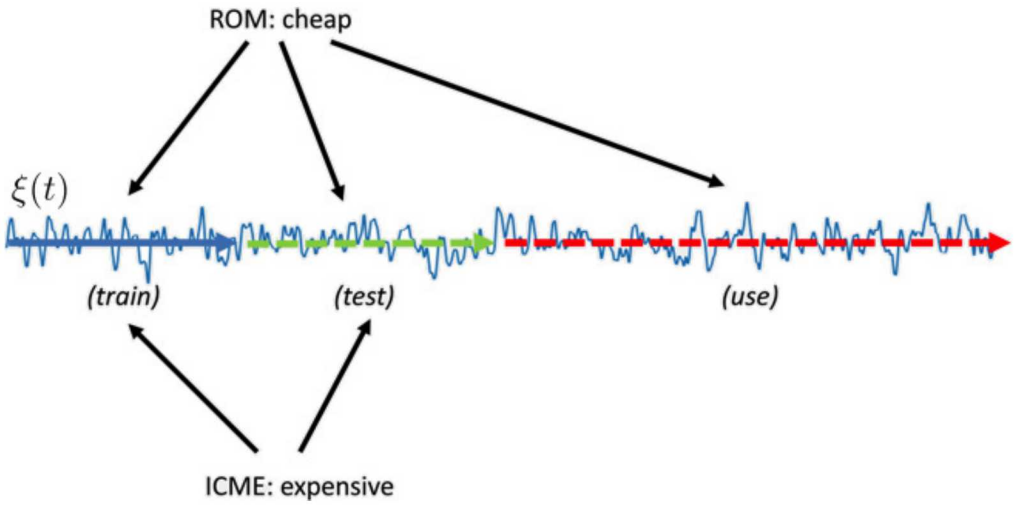


Figure 1: A schematic overview of the multiscale stochastic reduced-order model to accelerate uncertainty propagation in direct numerical simulation, showing the division of the time-series into two datasets, namely training and testing as in machine learning approaches. The ROM is trained using the first part of the time-series dataset, i.e. the training dataset. After the drift and diffusion coefficients are trained, the trained ROM is validated using the second part of the time-series dataset, i.e. the testing dataset. After the ROM is trained and validated, it can be deployed to propagate uncertainty beyond the time-scale limit of the ICME model.

4.2. Assumptions

Several assumptions are made in the ROM formulation. The assumptions are explained and justified as follows.

First, we assume that there is sufficiently enough data to train the ROM. The dataset can be obtained either experimentally through data acquisition, or computationally through running simulations repetitively on a high-performance computing platform. Some notable work to estimate the drift-diffusion parameters from experimental and computational time-series dataset include noisy electrical circuit [49], stochastic dynamics of metal cutting [50, 51, 52], meteorological data for sea surface wind [53, 54] to name a few. Interested readers are referred to the review paper of Friedrich et al. [2] for extensive multi-disciplinary applications across many scientific and engineering fields. Here, we applied and extended the method into the field of computational materials science with applications to microstructural evolution.

Second, we assume that the noise associated with QoIs are δ-correlated, in order to preserve the Markov property of the Langevin model. Furthermore, the noise is also assumed to be independent of the QoI $\xi(t)$. In practice, the noise is not strictly δ-correlated, which results in Markov-Einstein time τ_{ME}, such that for sampling intervals $\tau < \tau_{ME}$, the Markov property does not hold [45]. However, there has been proofs that the Markov assumption is a valid assumption, for example, in the field of fluid mechanics with small-scale turbulence [1]. Some recent efforts are also noted in adopting Fokker-Planck equation for non-Markovian process [55, 56].

Third, we assume that the drift and diffusion coefficients are a function of time, $D^{(n)}(x,t) = D^{(n)}(t)$. In literature, both time-independent [57] and time-dependent coefficients [58] have been studied. Generally speaking, the drift- and diffusion-coefficients do not have to be a temporal

function, but can be a spatio-temporal function, i.e. no restriction. This assumption is simply made for the computational convenience of the analytical calibration approach described in Section 4.3, but can simply be removed if the optimization approach in Section 4.3 is considered, as the optimization problem is considered as a black-box function and does not impose any restriction on the parameterization of the coefficients. This assumption only applies for the analytical approach, but not the optimization approach, during the calibration process for drift and diffusion coefficients.

Fourth, the proposed ROM is purely data-driven and does not have any underlying physical assumption. The purpose of this assumption is to retain the generalization of the proposed ROM to a wide range of ICME applications, without being restricted to a certain set of problems. However, it is possible to impose some physical conditions on the drift- and diffusion-coefficient, if it is desirable. The optional choice of imposing physical constraints depends on specific applications and is left to users. If the physical constraints are added, the ROM considered becomes a physics-constrained machine learning model, which is more restricted than the ROM considered in this paper.

Finally, we assume that only one QoI $\xi(t)$ is considered for a ROM. Theoretically, high-dimensional Fokker-Planck equations exist; practically, the Fokker-Planck equation is often solved in 2d, 3d [59], and 4d [60]. State-of-the-art mesh-based methods to solve Fokker-Planck equation through finite difference and finite element method are severely limited by the curse of dimensionality for more QoIs, which has a profound effect on the computing memory and speed. This assumption can be considered for computational convenience, opening up for future works to include more QoI in the same ROM. The assumptions for ROM construction is summarized as follows.

Assumption-1: There are sufficiently enough data, either experimentally or computationally, to train the ROM.

Assumption-2: The noise associated the QoIs are δ-correlated and independent of the QoIs, which preserves the Markov property.

Assumption-3: The drift and diffusion are functions of time, $D^{(n)}(x,t) = D^{(n)}(t)$, but not a function of x (for analytical calibration approach).

Assumption-4: The general ROM is data-driven, where there is no physical constraints on the drift and diffusion, although certain parameterization may be imposed depending on specific applications.

Assumption-5: For each ROM, only one QoI is considered.

4.3. Training drift and diffusion coefficients in the Fokker-Planck equation

The PDF for a QoI is numerically propagated along time, using the Fokker-Planck equation with calibrated coefficients and initial conditions. There are two important elements in constructing the stochastic ROM. First, the Fokker-Planck equation coefficients must be trained. Second, the initial conditions must be constructed with numerical stability consideration. The Fokker-Planck equation then can be solved using the calibrated coefficients and the initial conditions, and the QoI evolution along time can be predicted. During the training, the initial PDFs and evolution of PDFs associated with the QoIs are obtained by running the original materials simulation models, and the drift and

diffusion coefficients are calibrated based on the training PDFs. After calibration, the ROMs can be used to predict the PDFs of QoIs for longer periods of time, independently from the original material models. If necessary, additional calibration can be done at a later time period in order to ensure accurate predictions as the ROMs carry on. We propose two approaches to calibrate the stochastic ROM.

The first approach to analytically train or estimate coefficients is based on Corollaries 1 and 2 using linear regression. Based on Theorem 1, the coefficients can be estimated from the time-series dataset. However, in practice, direct application of Theorem 1 faces challenges from both spatio and temporal dimension. First, the number of QoI observations is often not sufficient in practice to approximate the central moments well enough along the spatial dimension. Second, the sampling time is often sparse along time dimension. As a result, numerical estimations based on derivatives to approximate coefficients, based on Theorem 1, are often noisy and oscillatory, creating numerical challenges to construct the ROM model. Observe that the QoI can be noisy in both temporal and spatial dimensions.

One approach to reduce the effect of noise is to perform homogenization for the spatial variable x in the coefficients and simplify the coefficients as functions of time t only, i.e. $D^{(n)}(x,t) = D^{(n)}(t)$. Excluding the spatial variable implies that the drift and diffusion coefficients are assumed to constant throughout the modeled spatial domain. This assumption is reasonable if the spatial domain is small. Materials distribution can be stable if a clear trend with respect to time is observed and can be anticipated in the future. The analytical approach built on Corollary 1 and Corollary 2 states that if the first two central moments are well constructed, then the drift and diffusion coefficients can be approximated by linear regression of mean and variance with respect to time, respectively. The numerical approximations converge to the central moment in probability by the weak law of large number [61].

A further challenge comes from the data used in calibration. It has been shown that the sampling rate of data is important in estimating the Fokker-Planck drift and diffusion coefficients from time-series data in previous studies (for example, Pienke et al. [62, 47, 48], Sura and Barsugli [63], Ragwitz et al. [64, 65], Honisch and Friedrich [45]). The data with low sampling rates are insufficient to estimate the coefficients accurately.

To circumvent the challenging problems posed by the first calibration approach, we also propose a second calibration approach to bypass the technical challenge of sampling rate by solving an inverse problem. Here, the coefficients are first parameterized, then optimized by minimizing the difference between the simulated and calibrated PDFs. The coefficients are trained and calibrated based on the minimization of a loss function. Compared to the first analytical approach described above, the second numerical approach is more robust.

The loss functions can be described as a distance between the PDFs from direct numerical simulations and the ROM predictions at a fixed time step, or at multiple time steps, where the predicted PDF with parameterized coefficients is compared with the simulated PDF as the training data. Mathematically, the loss functions can be expressed as either

$$l_1(x) = d\left(p^{(\tau)}_{\text{training}}(x), p^{(\tau)}_{\text{predicted}}(x)\right), \text{ at } t = \tau \text{ for a fixed time step } \tau, \tag{13}$$

or

$$l_2(x) = \sum_{j=1}^{n} w_j d\left(p_{\text{training}}^{(\tau_j)}(x), p_{\text{predicted}}^{(\tau_j)}(x)\right), \text{ at } t = \tau_j \text{ for multiple time steps } (\tau_j)_{j=1}^{n} \tag{14}$$

where $d(\cdot,\cdot)$ is the distance between two PDFs. The distances $d(\cdot,\cdot)$ can be defined in different ways, such as l_p and L^P norms, Kullback-Leibler divergence, Wasserstein distance, or others.

4.4. Regularization for initial conditions

After the ROMs are calibrated, the evolution of PDFs for QoIs can be efficiently simulated with much longer time steps. The numerical stability of the ordinary differential equations however is sensitive to the initial PDFs. 'Noisy' empirical PDFs of QoIs from direct numerical simulations need to be processed with regularization, before they are used as the initial conditions in solving the Fokker-Planck equations.

Here the ridge regression method, which falls under the class of Tikhonov regularization [66], is applied to smoothen out the empirical initial PDF for solving the forward Fokker-Planck equation. The goal of regularization is to seek for an approximated PDF $\hat{f}(x)$ in a bounded domain $x \in \mathbb{R}$ that minimizes the penalized least squares function

$$Q(\hat{f}) = \int_{x_1}^{x_N} \left|\hat{f}(x) - f(x)\right|^2 dx + \lambda \int_{x_1}^{x_N} \left(\frac{\partial^d \hat{f}(x)}{\partial x^d}\right)^2, \tag{15}$$

where λ is the regularization parameter, and d is the order of derivative.

When discretized, the objective function Q can be expressed as

$$Q = (\mathbf{M}\hat{\mathbf{f}} - \mathbf{f})^{\mathbf{T}}\mathbf{F}^{-\mathbf{2}}\mathbf{B}(\mathbf{M}\hat{\mathbf{f}} - \mathbf{f}) + \lambda(\mathbf{E}\hat{\mathbf{f}})^{\mathbf{T}}\tilde{\mathbf{B}}(\mathbf{E}\hat{\mathbf{f}}), \tag{16}$$

where $\boldsymbol{E}$ is the derivative matrix of any order, $\mathbf{M}$ is the mapping matrix, $\mathbf{B}$ is the midpoint rule integration matrix, $\tilde{\mathbf{B}}$ is a subset of $\mathbf{B}$, and $\mathbf{F} = \text{diag}(\mathbf{f})$ is the observation matrix in diagonal form.

Setting $\mathbf{M} = \mathbf{F}^{-\mathbf{2}}\mathbf{B} = \tilde{\mathbf{B}} = \mathbf{I}$ and solving $\partial Q/\partial f = 0$, we obtain the simplest form of smoothing by regularization, as

$$\hat{\mathbf{f}} = (\mathbf{I} + \lambda\mathbf{E}^{\mathbf{T}}\mathbf{E})^{-\mathbf{1}}\mathbf{f}. \tag{17}$$

4.5. Finite difference Fokker-Planck equation solver

Numerically the Fokker-Planck equation is commonly solved in two ways, finite element method and finite difference method. Both of them suffer the curse of dimensionality for large problems with high-dimensional state space. However, in the scope of this paper, only 1D stochastic process is concerned. Finite difference method is applied here.

The numerical implementation of finite difference method is developed based on the algorithm of Hassan et al. [67], which calculates derivatives of any degree with any arbitrary order of accuracy over a uniform grid. The Fokker-Planck equation in 4 is discretized and implemented with a matrix form as

$$\dot{\boldsymbol{f}}(x,t) = -\boldsymbol{E}^{(1)}\left[D^{(1)}(x,t)\boldsymbol{f}(x,t)\right] + \boldsymbol{E}^{(2)}\left[D^{(2)}(x,t)\boldsymbol{f}(x,t)\right], \tag{18}$$

with the initial condition is $\boldsymbol{f}(x, t = \tau_0)$.

After discretization in spatio-temporal dimensions, Equation 18 can be numerically solved explicitly using the Runge-Kutta method, or implicitly using Crank-Nicolson method [68].

5. Applications and demonstrations

In this section, the proposed ROM is demonstrated using three examples: kMC simulation in Section 5.1, PF simulation in Section 5.2, and MD simulation in Section 5.3. In the kMC example, the grain growth is simulated with a hybrid Potts-phase field model. The selected QoI is the grain area. In the PF example, the evolution of Fe-Cr microstructures and phases is simulated. The QoI is the chord-length. In the MD example, a simple liquid argon system is simulated, and the selected QoIs are the total mean-displacements and enthalpy of the simulation cell. In these examples, the ICME models are ran for a period of time. Then, the QoI' PDFs are obtained by post-processing the simulations. A beginning portion with respect to time of the PDFs collection is considered as the training dataset, as illustrated in Figure 1, whereas the last portion of this PDFs collection is considered as the testing dataset. The comparison between the evolving PDF of the Fokker-Planck equation after calibrated using the training dataset and the last PDF in the testing dataset is a reasonable measure on the numerical performance of the proposed stochastic ROM method.

5.1. Kinetic Monte Carlo simulation: hybrid Potts-phase field simulation for grain growth

In this example, the hybrid Potts-phase field model from Homer et al. [69] based on kinetic Monte Carlo SPPARKS framework [70] is used to investigate the evolution of the grain area during the grain growth. The hybrid Potts-phase field model is applied on a simple two-component, two-phase system, where the bulk free energy of the system is described as [69]

$$E_v(q, C) = \lambda[(C - C_1)^2 + (C_2 - C)^2] + a(C - C_3)^2 q_\alpha + \alpha(C_4 - C)^2 q_\beta, \tag{19}$$

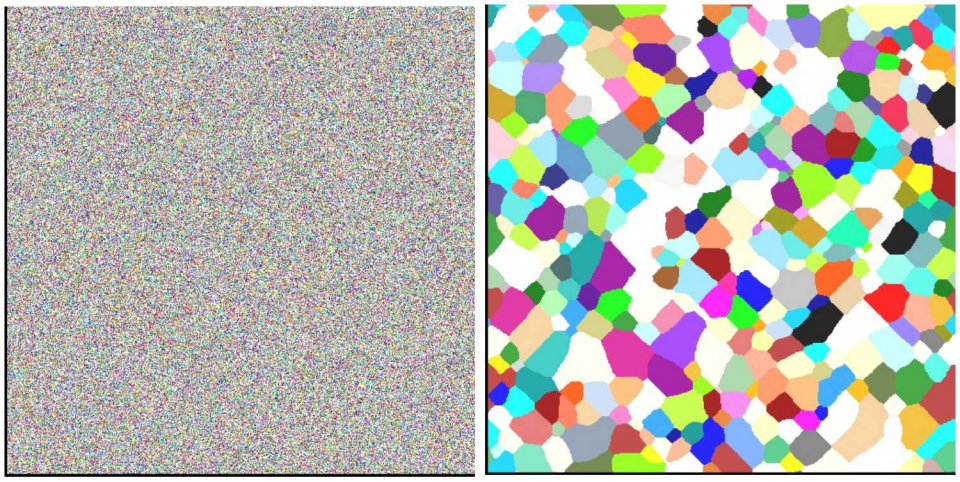
where $\lambda = 0.3, C_1 = 0.25, C_2 = 0.75, C_3 = 0.05, C_4 = 0.95$, and $\alpha = 0.5$. A computational domain of 5000 pixel $\times$ 5000 pixel is used to perform the 2D grain growth kMC simulation. The kMC simulation is ran for 20,000 Monte Carlo steps (mcs), where 40 Monte Carlo events are observed, with the last microstructure is obtained at 16,681.1 mcs.

The drift and diffusion coefficients are calibrated using the first analytical approach based on Corollaries 1 and 2. The initial and training PDFs are constructed using the kernel density estimation method with the normal kernel distribution. The selected bandwidth is optimal for the normal kernel density [71]. The initial PDF is constructed and regularized using the Tikhinov regularization as described in Section 4.4 to reduce the chance of divergent Fokker-Planck solution. In the total of 40 Monte Carlo events, the first Monte Carlo 34 events, which describes the grain evolution from 0 mcs to 4641.62 mcs, are used as the training dataset, while the rest, which are up to 16,681.1 mcs, are used as the testing dataset.

Figure 2 shows the evolution of microstructure using kMC. In parallel, Figure 3 shows the evolution of the QoI using the Fokker-Planck equation at the same steps with Figure 2, where the QoI is grain area based on the kMC simulation results. The calibrated density is the last PDF used to train the stochastic ROM. The final density denotes the last PDF, obtained from direct simulations, to evaluate the performance of the trained stochastic ROM. The final density is not used for training the stochastic ROM.

In Figure 3a, the initial PDF of grain area is peaked at the size of approximately 5000 pixel^2, corresponding to Figure 2a, as the grains are fairly small and uniform. This is shown by the small

variance of the QoI in Figure 2a. As the simulation continues, the grain grows larger, and the variance of the QoI PDF increases accordingly. Figure 2d and Figure 3d present the microstructure and its corresponding QoI's PDF after the Fokker-Planck coefficients have been calibrated, respectively. In Figure 3d, the testing PDF and evolved Fokker-Planck PDF after calibration agree very well at the latest testing time. This demonstrates if the Fokker-Planck coefficients are well-trained, a prediction about the evolution of the microstructural descriptor using the trained ROM can be made with a good level of accuracy.



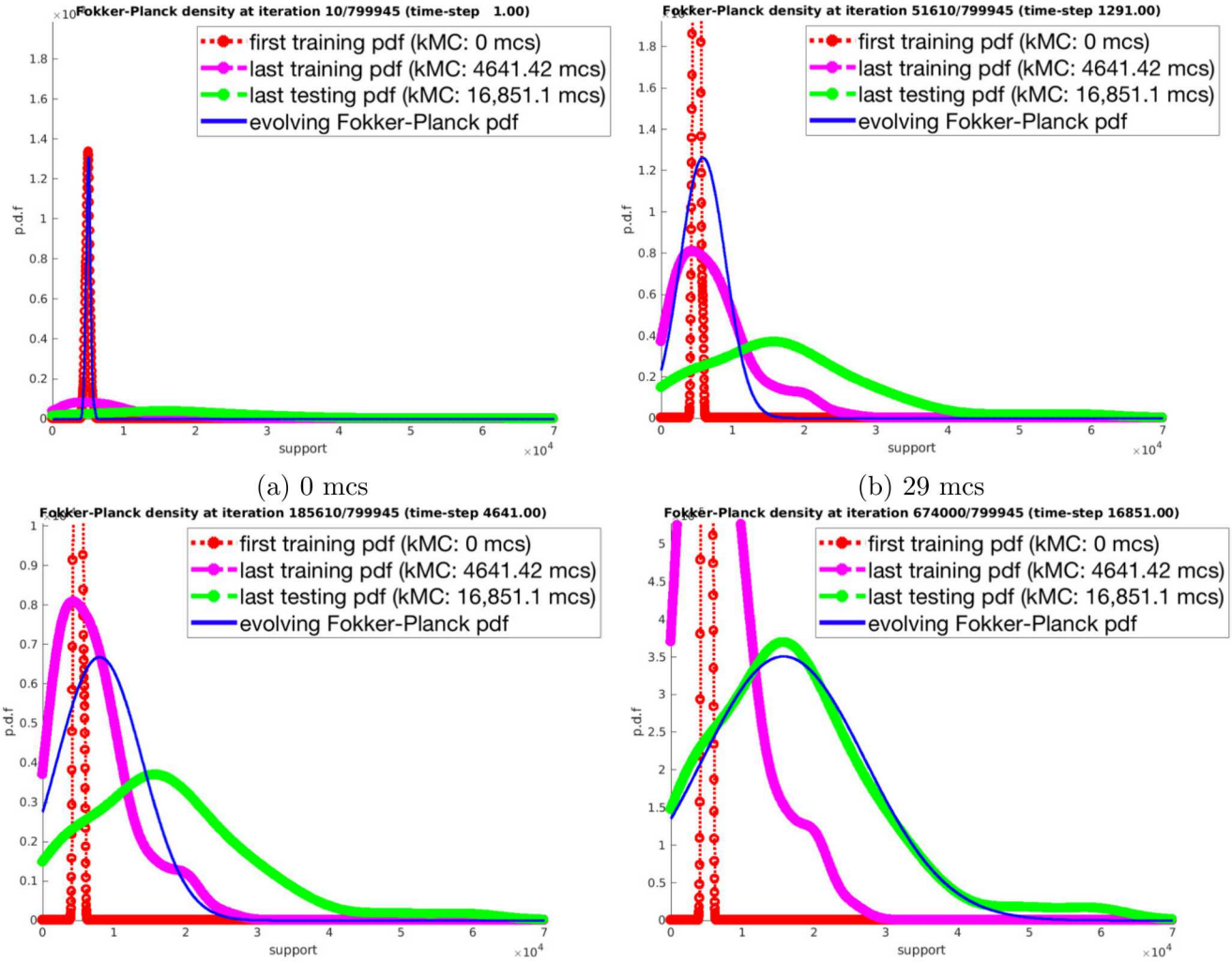
(a) Initialization: microstructure at 0 mcs

(b) 29^{th} microstructure at 1291.62 mcs

(c) 34^{th} microstructure at 4641.62 mcs

(d) 39^{th} microstructure at 16,681.1 mcs

Figure 2: Microstructural evolution of grain growth in kMC simulation.



(a) 0 mcs

(b) 29 mcs

(c) 34 mcs

(d) 39 mcs

Figure 3: Evolution of grain area distribution. The first training PDF is at 0 mcs, the last training PDF is at 4641.42 mcs, and the last testing PDF is at 16,851.1 mcs. (Readers are referred to online version to visualize different colors. Red curve denotes the first training PDF, magenta curve denotes the last training PDF, green curve denotes the testing PDF, and the blue curve denotes the predicted PDF by the Fokker-Planck equation after calibration.)

5.2. Phase field simulation: Spinodal decomposition

In this example, the Fe-Cr microstructural evolution using the PF simulation in the MOOSE framework [72] is used to demonstrate the approach. In this example, the QoI is the chord-length distribution, which is another statistical microstructural descriptor. The training and initial PDFs are also constructed using the kernel density estimation method with the normal kernel distribution. The optimal bandwidth is selected for the normal kernel density [71]. The initial PDF is regularized using the Tikhinov regularization as described in Section 4.4 to reduce the probability of divergence Fokker-Planck solution.

The training dataset is the PDFs collection from 0 steps to 1700 steps, where the rest, which are the PDFs collection from 1700 steps to 2405 steps, is the testing dataset. Due to the noise and instability observed at the beginning of the simulation, the training dataset is truncated from 1400 steps to 1700 steps to exclude extreme variations at the beginning of the PF simulation. This variation is typically observed in many dynamic ICME models, including MD and PF simulations.

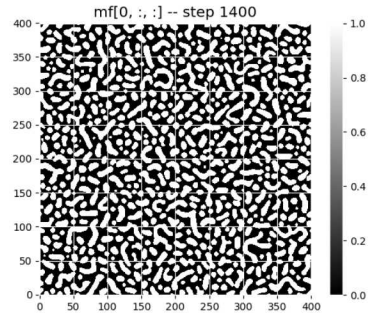
Figure 4 shows the microstructural evolution of Fe-Cr spinodal decomposition simulations on 25 nm $\times$ 25 nm at 500°C over a period of 7 days (604800s) in physical time. The system is modeled using the Cahn-Hilliard equation with no external energy sources. The initial concentration of Cr is randomly generated within the interval [44.774%, 48.774%] with the expectation of 46.774%. The coarsening effect is observed, and the clusters slowly expand as the simulation advances.

The coefficients of the Fokker-Planck equation are calibrated using the Bayesian optimization method [73, 74, 75, 77]. Here, the drift and diffusion coefficients are parameterized, and the batch-parallel Bayesian optimization is applied to minimize the Kullback-Leibler divergence between the training PDF and the predicted Fokker-Planck PDF, as $\mathrm{KL}(p_{\text{training}}||p_{\text{predicted}})$, where the $p_{\text{predicted}}$ is obtained from solving the forward Fokker-Planck equation with certain coefficients.

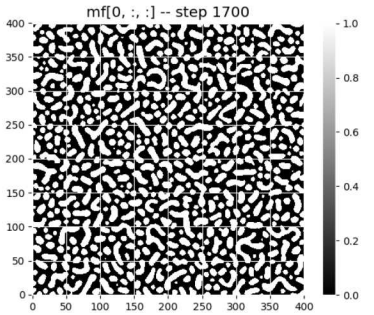
Figure 5 presents the evolution of calibrated Fokker-Planck equation to capture the evolution of QoI. Figure 4d shows the comparison between the PDF obtained by calibrated and trained Fokker-Planck equation and the testing PDF from the ICME model, which is the PF simulation in this case.

5.3. Molecular dynamics simulation: Equilibrium liquid Argon

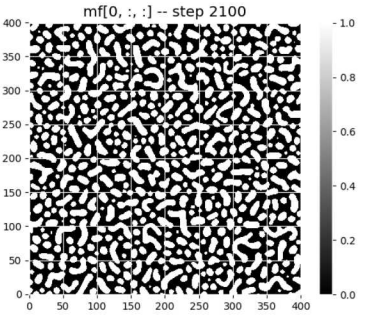
In this example, MD simulation of liquid Argon at 85K is performed using LAMMPS [78] to assess the total mean square displacement and enthalpy of the simulated system. The system consists of 4000 atoms, where the interatomic potential is described by Lennard-Jones model with uncertain well-depth ε and well-location σ. Different ε and σ for Argon have been used in the literature, for example, McGaughey et al. [79], Borgelt et al [80], Dawid et al. [81], Laasonen et al. [82], Reith et al. [83], Griebel et al. [84]. The uncertain ε and σ here are modeled with truncated normal distributions. The mean μ, the standard deviation σ, the support lower and upper bounds are (0.2383,0.0667,0.2376,0.2390) for ε and (3.4000,0.6670,3.3000,3.5000) for σ, respectively. The microcanonical ensemble (NVE) is used, where a Langevin thermostat is also used to coupled with the system. The Langevin thermostat, which has a random noise generator [85], can be thought of a source for aleatory uncertainty. 25,062 equilibrium simulations are performed, and the QoIs are analyzed using log files of the simulation. The sampling time is 50 fs, the time step is 1 fs, and the total simulation time is 20 ps. The collection of PDFs from 0 ps to 160 ps is used as the training



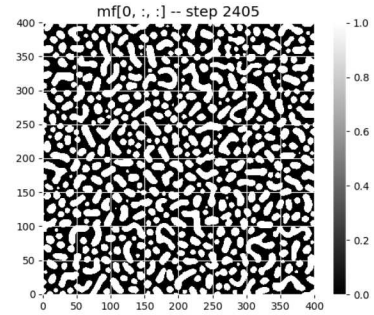
(a) 1400τ



(b) 1700τ

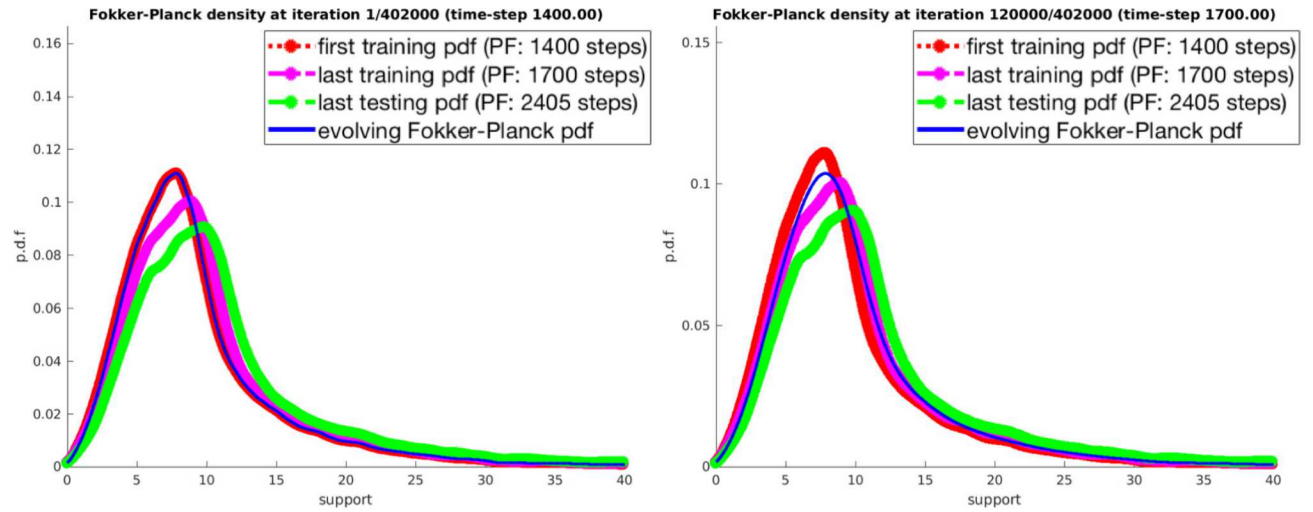


(c) 2100τ



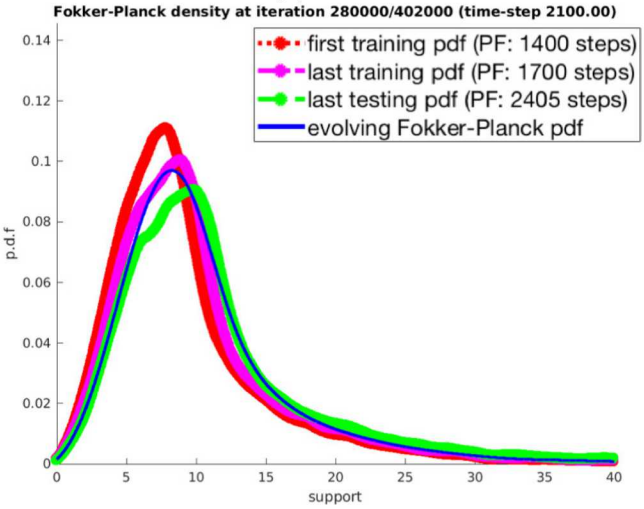
(d) 2405τ

Figure 4: Microstructural evolution in PF spinodal decomposition simulation. Two phases exist in this system: the Fe-rich and Cr-rich phases. The Fe-rich phase is plotted as white, wheresa the Cr-rich phase is plotted as black.

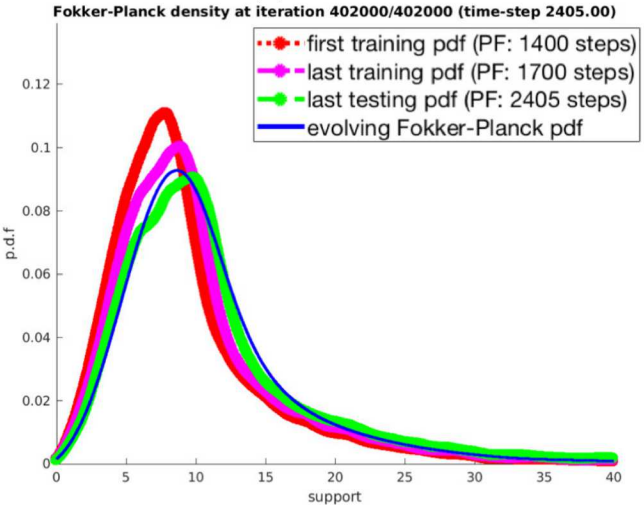


(a) 1400τ

(b) 1700τ



(c) 2100τ



(d) 2405τ

Figure 5: Evolution of chord-length distributions. The first training PDF is at 1400 steps, the last training PDF is at 1700 steps, and the last testing PDF is at 2405 steps. (Readers are referred to online version to visualize different colors. Red curve denotes the first training PDF, magenta curve denotes the last training PDF, green curve denotes the testing PDF, and the blue curve denotes the predicted PDF by the Fokker-Planck equation after calibration.)

dataset, whereas the collection of PDFs from 160 ps to 200 ps is used as the testing dataset. Due to the instability of the MD simulation, only a part of the training dataset from 140 ps to 160 ps is used to exclude the noise.

Monte-Carlo sampling is used to assess the *a posteriori* distribution of the QoIs. The training and initial PDFs are reconstructed using kernel density estimation method with the normal kernel distribution. The selected bandwidth is optimal for the normal kernel density [71]. The initial PDF for total mean-square displacement is regularized using the Tikhinov regularization as described in Section 4.4 to avoid divergent solutions in solving the Fokker-Planck equation.

Figure 6 and Figure 7 show the evolution of the QoIs' PDFs in different snapshots at 140, 160, 180, and 200 ps. In Figure 7, all the PDFs are fitted to normal distributions instead of approximating by the kernel density estimation method. The Fokker-Planck coefficients for the total mean-square displacement are trained by minimizing the Kullback-Leibler divergence, whereas the coefficients for enthalpy are trained using Corollary 1 and Corollary 2. The comparison between the testing PDF and evolved Fokker-Planck PDF shows a fairly good agreement after the Fokker-Planck coefficients are calibrated.

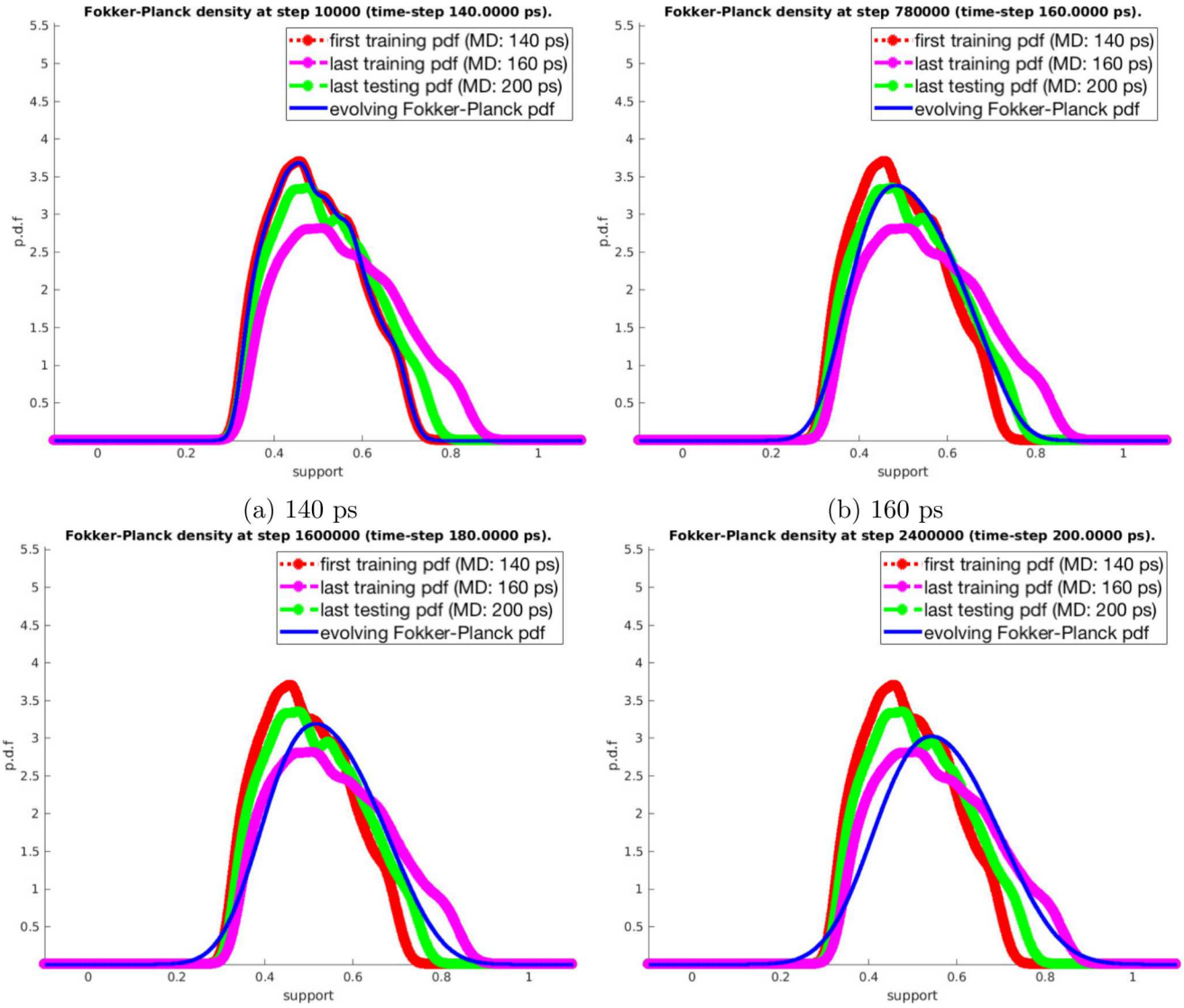
6. Discussions

We compare the computational cost on a AMD A10-6700 CPU at 3.7GHz on Ubuntu 18.04 platform between optimized C/C++ packages for ICME models and a preliminary non-optimized implementation on MATLAB, where the results are tabulated in Table 1. Efficient and optimized C++ implementation of the proposed ROM in Trilinos [86, 87] or PETSc [88, 89] would increase the speedup factor significantly. The speedup depends on various factors, such as the computational time of the ICME model (which varies depending on the fidelity of the model), how many times the simulation repeats, the implementation of ROM (the time-step used in the integrator, as well as the robustness of the partial differential equation linear solver for the ROM). Except for the kMC and PF, in which only one large-scale simulation is performed, for MD, 25,062 small-scale simulations are performed, with each simulation costs around 0.8254 hr, because each MD simulation can only sample one trajectory. The proposed ROM is particularly competitive with a large speedup when ICME models are computationally expensive, such as high-fidelity large-scale simulations.

Table 1: Comparison of computation cost between ICME models and ROM.

	ICME packages	ICME comp.	ROM comp.	Speedup
kMC	SPPARKS	76.3486 hr	0.1375 hr	555.2625x
PF	MOOSE	97.8667 hr	1.2189 hr	80.2909x
MD	LAMMPS	20,686.1748 hr	65.8581 hr	314.1022x

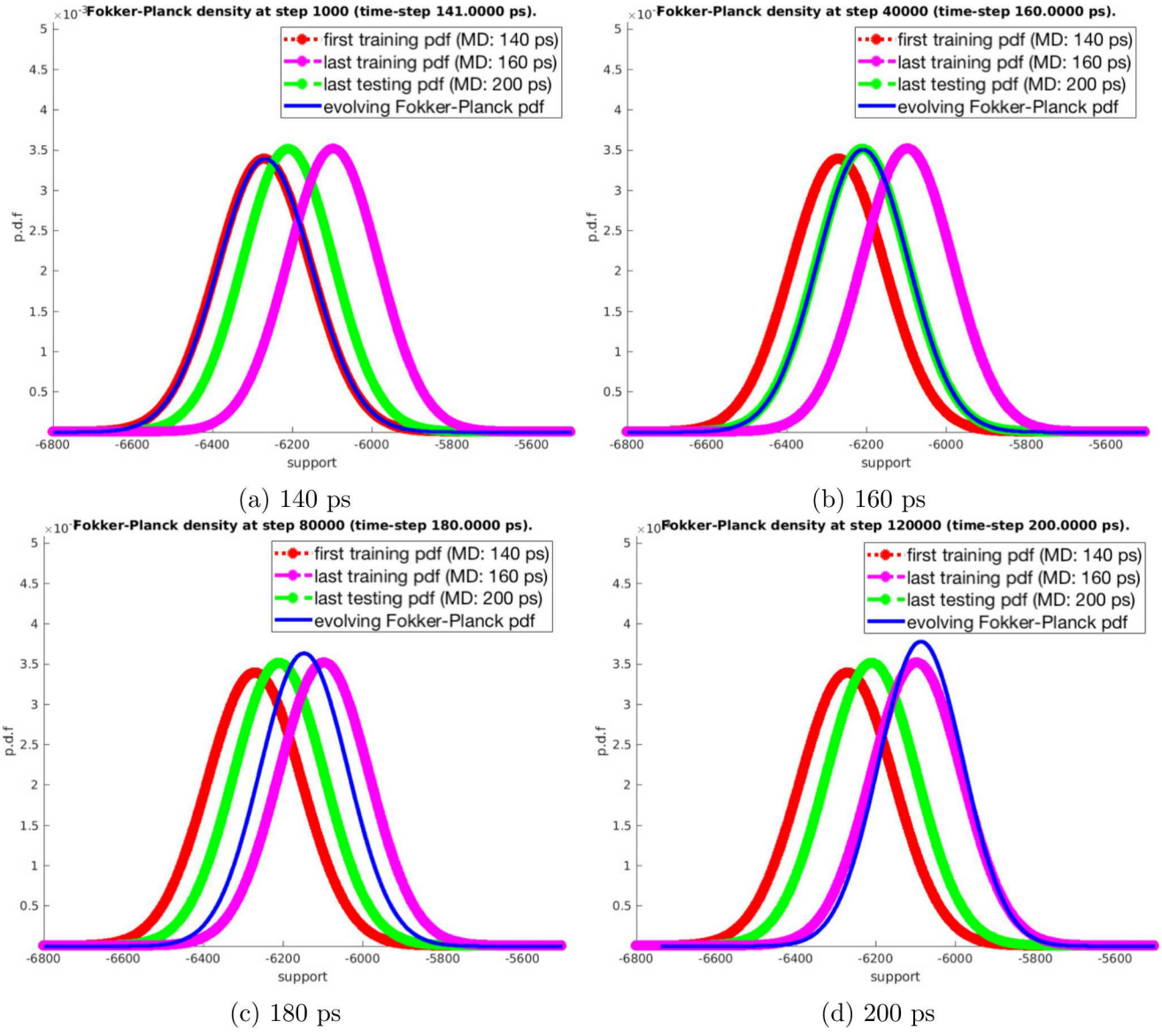
In this paper, we present a multiscale stochastic ROM that accelerates the uncertainty propagation in materials modeling of system dynamics, by modeling the evolution of uncertain QoIs using the Fokker-Planck equation. The proposed method is demonstrated by both drift-dominated (as in the cases of MD and kMC) and diffusion-dominated (as in the case of PF) examples. It is shown that if the Fokker-Planck equation coefficients are appropriately parameterized and well calibrated,



(a) 140 ps (b) 160 ps

(c) 180 ps (d) 200 ps

Figure 6: Evolution of total mean-square displacement distributions. The first training PDF is at 140 ps, the last training PDF is at 160 ps, and the last testing PDF is at 200 ps. (Readers are referred to online version to visualize different colors. Red curve denotes the first training PDF, green curve denotes the training PDF, magenta curve denotes the testing PDF, and the blue curve denotes the predicted PDF by the Fokker-Planck equation after calibration.)



(a) 140 ps (b) 160 ps

(c) 180 ps (d) 200 ps

Figure 7: Evolution of enthalpy in MD simulation. The first training PDF is at 140 ps, the last training PDF is at 160 ps, and the last testing PDF is at 200 ps. (Readers are referred to online version to visualize different colors. Red curve denotes the first training PDF, green curve denotes the training PDF, magenta curve denotes the testing PDF, and the blue curve denotes the predicted PDF by the Fokker-Planck equation after calibration.)

the proposed method has a predictive capacity to estimate the QoI distributions for longer time scales, without using the computationally expensive material simulations.

The diffusion process modeling by the Fokker-Planck equation is used as a stochastic ROM in this paper. If there is no diffusion, then the diffusion coefficient becomes zero, while the drift coefficient is non-zero. We assume that the QoI can be modelled by the diffusion process. However, this is only true for certain applications, such as those driven by diffusion phenomena in materials science, where the materials behaviors are experimentally justified.

In addition, we also assume that the drift and diffusion coefficients are only a function of time, and not a function of QoI. This assumption conveniently simplifies the estimation of the drift and diffusion coefficients, yet also imposes the same effect regardless of the QoI magnitudes, posing a drawback of the proposed method and a potential future study. Interested readers are referred to the works of Friedrich et al. [2, 45, 47, 49, 65] for various applications of stochastic method in time-series and estimation the drift and diffusion coefficients of Fokker-Planck equations in low- and high-sampling rates.

The multiscale efficiency is mainly based on the difference between the time-scale of solving the Fokker-Planck equation and the time-scale of solving the ICME models. The multiscale efficiency depends on the cost of solving the Fokker-Planck equation, as well as the cost of simulating the ICME models. The cost of solving the Fokker-Planck equation can be significantly improved by an implicit time-integrator, as well as a preconditioner for the Fokker-Planck solver. The time-scale of solving the Fokker-Planck equation depends on the scale of ICME models, which can be large scale in both length-scale and time-scale. Because these two time-scales (one of the ICME models and one of the Fokker-Planck equation) are completely independent, the benefits of using the stochastic ROM depend on many factors. In general, the benefits are maximized if the ICME models are very computationally expensive, and if the Fokker-Planck equation can be efficiently solved. The improvement of Fokker-Planck solver is posed as open questions for further research.

In this paper, only one QoI is considered in a ROM. Building ROMs for coupled or correlated QoIs requires the high-dimensional Fokker-Planck equation, which will be a topic of future study. If the QoIs are structural descriptors, their predicted values can also be used to reconstruct the microstructure. Examples of statistical and deterministic microstructural descriptors are discussed in Torquato et al. [90] and Groeber et al. [91, 92], Chen et al. [93, 94]. The statistical descriptors tend to outnumber the deterministic descriptors, due to the random nature of materials. The proposed framework in this paper can potentially be used to predict the evolution of microstructures for a much longer time scale than the direct simulations can achieve.

The Kramers-Moyal expansion allows one to propagate higher-order moments in theory, but in practice, estimations of the Kramers-Moyal coefficients can also be approached using optimization methods. Furthermore, higher-order derivative terms make the partial differential equation harder to numerically solve for the forward Kramers-Moyal expansion. The proposed method is extensible in two directions. The former extension includes more variables in a high-dimensional Fokker-Planck equation, whereas the later extension would capture the QoIs evolution with smaller approximation errors by including higher-order terms for the same QoI. A caveat of the later extension is that with higher-order derivatives, it is likely that the numerical solution of the forward Kramers-Moyal expansion would diverge at some points, and some numerical treatments are required in order to

obtain a convergent solution. A robust implicit integration solver may be required to stabilize in solving higher-order Kramer-Moyal expansion. Tikhonov regularization can also be applied to smooth out the PDFs.

7. Conclusion

In this paper, a multiscale stochastic ROM method is proposed to solve the time-scale issue of uncertainty propagation in materials modeling. The ROM coefficients are trained either analytically or numerically, so that the evolution of QoIs can be accurately captured using the ROMs. The Ornstein-Uhlenbeck stochastic process is modeled using 1D generalized Langevin equation. With the formulation of Stratonovich calculus, the stochastic variable can be modeled using the Fokker-Planck equation.

Three examples are used to demonstrate the multiscale stochastic ROM framework, including kMC, PF, and MD, where the statistical microstructural descriptors are the QoIs. The results show a good agreement between the prediction from the trained ROMs and the direct simulations, when the ROMs are parameterized and calibrated appropriately.

Acknowledgments

The research was supported by NSF under grant number CMMI-1306996 and the George W. Woodruff Faculty Fellowship. The research cyberinfrastructure resources and services provided by the Partnership for an Advanced Computing Environment (PACE) at the Georgia Institute of Technology are appreciated.

The views expressed in the article do not necessarily represent the views of the U.S. Department of Energy or the United States Government. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525.

Appendix A. Analytical solutions of Fokker-Planck equation

Here we present two families of Gaussian distribution, with three simple analytical examples that can capture a wide range of phenonmena with increasing complexity. These examples are the analytical solution of the one-dimensional Fokker-Planck equation as in Equation 4.

The first example is a Gaussian probability distribution function with no drift and constant diffusion parameter,

$$f_1(x,t) = \frac{1}{\sqrt{4\pi Dt}} e^{-\frac{x^2}{4Dt}}. \tag{A.1}$$

It is easy to see that for $f_1(x,t)$, $\mathbb{E}[\xi(t)] = 0$, $D^{(1)} = \dfrac{\partial \mathbb{E}[\xi(t)]}{\partial t} = 0$. $\mathrm{Var}[\xi(t)] = 2Dt$, $D^{(2)} = \dfrac{1}{2}\dfrac{\partial \mathrm{Var}[\xi(t)]}{\partial t} = D$.

The second example is a Gaussian probability distribution function with constant drift, where the mean is moving with a constant velocity and no diffusion,

$$f_2(x,t) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu t)^2}{2\sigma^2}}. \tag{A.2}$$

In $f_2(x,t)$ case, $\mathbb{E}[\xi(t)] = \mu t$, $D^{(1)} = \dfrac{\partial \mathbb{E}[\xi(t)]}{\partial t} = \mu$. $\mathrm{Var}[\xi(t)] = \sigma^2$, $D^{(2)} = \dfrac{1}{2}\dfrac{\partial \mathrm{Var}[\xi(t)]}{\partial t} = 0$.

The third example, which is the most general among these examples, is a Gaussian probability distribution function with constant drift, where the mean is moving with a constant velocity and constant diffusion, where the variance increases linearly with respect to time,

$$f_3(x,t) = \frac{1}{\sqrt{4\pi D t}} e^{-\frac{(x-\mu t)^2}{4\pi D t}}. \tag{A.3}$$

For $f_3(x,t)$, $\mathbb{E}[\xi(t)] = \mu t$, $D^{(1)} = \dfrac{\partial \mathbb{E}[\xi(t)]}{\partial t} = \mu$. $\mathrm{Var}[\xi(t)] = 2Dt$, $D^{(2)} = \dfrac{1}{2}\dfrac{\partial \mathrm{Var}[\xi(t)]}{\partial t} = D$.

References

[1] C. Renner, J. Peinke, R. Friedrich, Experimental indications for Markov properties of small-scale turbulence, Journal of Fluid Mechanics 433 (2001) 383–409.

[2] R. Friedrich, J. Peinke, M. Sahimi, M. R. R. Tabar, Approaching complexity by stochastic methods: From biological systems to turbulence, Physics Reports 506 (5) (2011) 87–162.

[3] M. Grigoriu, Reduced order models for random functions. application to stochastic problems, Applied Mathematical Modelling 33 (1) (2009) 161–175.

[4] S. Sarkar, J. E. Warner, W. Aquino, M. D. Grigoriu, Stochastic reduced order models for uncertainty quantification of intergranular corrosion rates, Corrosion Science 80 (2014) 257–268.

[5] M. P. Mignolet, C. Soize, Stochastic reduced order models for uncertain geometrically nonlinear dynamical systems, Computer Methods in Applied Mechanics and Engineering 197 (45-48) (2008) 3951–3963.

[6] R. G. Ghanem, P. D. Spanos, Stochastic finite elements: a spectral approach, Courier Corporation, 2003.

[7] T. Goudon, L. Monasse, Fokker-Planck approach of Ostwald ripening: simulation of a modified Lifschitz-Slyozov-Wagner system with a diffusive correction.

[8] B. Ganapathysubramanian, N. Zabaras, Modeling diffusion in random heterogeneous media: Data-driven models, stochastic collocation and the variational multiscale method, Journal of Computational Physics 226 (1) (2007) 326–353.

[9] S. Bhattacharjee, K. Matouš, A nonlinear manifold-based reduced order model for multiscale analysis of heterogeneous hyperelastic materials, Journal of Computational Physics 313 (2016) 635–653.

[10] M. I. Latypov, S. R. Kalidindi, Data-driven reduced order models for effective yield strength and partitioning of strain in multiphase materials, Journal of Computational Physics 346 (2017) 242–261.

[11] M. J. Gander, 50 years of time parallel time integration, in: Multiple Shooting and Time Domain Decomposition Methods, Springer, 2015, pp. 69–113.

[12] J. Nievergelt, Parallel methods for integrating ordinary differential equations, Communications of the ACM 7 (12) (1964) 731–733.

[13] P. Chartier, B. Philippe, A parallel shooting technique for solving dissipative ODE's, Computing 51 (3-4) (1993) 209–236.

[14] P. Saha, J. Stadel, S. Tremaine, A parallel integration method for solar system dynamics, arXiv preprint astro-ph/9605016.

[15] Y. Maday, G. Turinici, A parareal in time procedure for the control of partial differential equations, Comptes Rendus Mathematique 335 (4) (2002) 387–392.

[16] G. Bal, On the convergence and the stability of the parareal algorithm to solve partial differential equations, in: Domain decomposition methods in science and engineering, Springer, 2005, pp. 425–432.

[17] E. Lelarasmee, The waveform relaxation method for time domain analysis of large scale integrated circuits: Theory and applications, Electronics Research Laboratory, College of Engineering, University of California, 1982.

[18] M. J. Gander, Overlapping Schwarz for linear and nonlinear parabolic problems.

[19] M. J. Gander, L. Halpern, F. Nataf, Optimal convergence for overlapping and non-overlapping Schwarz waveform relaxation.

[20] W. Hackbusch, Parabolic multi-grid methods, in: Proc. of the sixth int'l. symposium on Computing methods in applied sciences and engineering, VI, North-Holland Publishing Co., 1985, pp. 189–197.

[21] C. Lubich, A. Ostermann, Multi-grid dynamic iteration for parabolic equations, BIT Numerical Mathematics 27 (2) (1987) 216–234.

[22] G. Horton, S. Vandewalle, A space-time multigrid method for parabolic partial differential equations, SIAM Journal on Scientific Computing 16 (4) (1995) 848–864.

[23] M. Emmett, M. L. Minion, et al., Toward an efficient parallel in time method for partial differential equations, Communications in Applied Mathematics and Computational Science 7 (1) (2012) 105–132.

[24] M. Neumüller, Space-time methods: fast solvers and applications, Ph.D. thesis, University of Graz (2013).

[25] W. L. Miranker, W. Liniger, Parallel methods for the numerical integration of ordinary differential equations, Mathematics of Computation 21 (99) (1967) 303–320.

[26] A. Axelsson, J. G. Verwer, Boundary value techniques for initial value problems in ordinary differential equations, Mathematics of Computation 45 (171) (1985) 153–171.

[27] D. E. Womble, A time-stepping algorithm for parallel computers, SIAM Journal on Scientific and Statistical Computing 11 (5) (1990) 824–837.

[28] P. Worley, Parallelizing across time when solving time-dependent partial differential equations, in: Proc. 5th SIAM Conf. on Parallel Processing for Scientific Computing, D. Sorensen, ed., SIAM, 1991.

[29] D. Sheen, I. H. Sloan, V. Thomée, A parallel method for time discretization of parabolic equations based on Laplace transformation and quadrature, IMA Journal of Numerical Analysis 23 (2) (2003) 269–299.

[30] Y. Maday, E. M. Rønquist, Parallelization in time through tensor-product space–time solvers, Comptes Rendus Mathematique 346 (1-2) (2008) 113–118.

[31] A. J. Christlieb, C. B. Macdonald, B. W. Ong, Parallel high-order integrators, SIAM Journal on Scientific Computing 32 (2) (2010) 818–835.

[32] S. Güttel, A parallel overlapping time-domain decomposition method for odes, in: Domain decomposition methods in science and engineering XX, Springer, 2013, pp. 459–466.

[33] T. Arbogast, G. Pencheva, M. F. Wheeler, I. Yotov, A multiscale mortar mixed finite element method, Multiscale Modeling & Simulation 6 (1) (2007) 319–346.

[34] B. Ganis, I. Yotov, Implementation of a mortar mixed finite element method using a multiscale flux basis, Computer Methods in Applied Mechanics and Engineering 198 (49-52) (2009) 3989–3998.

[35] W. E, B. Engquist, et al., The heterognous multiscale methods, Communications in Mathematical Sciences 1 (1) (2003) 87–132.

[36] A. Abdulle, W. E, B. Engquist, E. Vanden-Eijnden, The heterogeneous multiscale method, Acta Numerica 21 (2012) 1–87.

[37] T. J. Hughes, L. Mazzei, K. E. Jansen, Large eddy simulation and the variational multiscale method, Computing and Visualization in Science 3 (1-2) (2000) 47–59.

[38] T. J. Hughes, A. A. Oberai, L. Mazzei, Large eddy simulation of turbulent channel flows by the variational multiscale method, Physics of fluids 13 (6) (2001) 1784–1799.

[39] Y. Bazilevs, I. Akkerman, Large eddy simulation of turbulent taylor–couette flow using isogeometric analysis and the residual-based variational multiscale method, Journal of Computational Physics 229 (9) (2010) 3402–3414.

[40] R. E. Miller, E. B. Tadmor, A unified framework and performance benchmark of fourteen multiscale atomistic/continuum coupling methods, Modelling and Simulation in Materials Science and Engineering 17 (5) (2009) 053001.

[41] G. J. Wagner, W. K. Liu, Coupling of atomistic and continuum simulations using a bridging scale decomposition, Journal of Computational Physics 190 (1) (2003) 249–274.

[42] W. A. Curtin, R. E. Miller, Atomistic/continuum coupling in computational materials science, Modelling and simulation in materials science and engineering 11 (3) (2003) R33.

[43] H. Risken, The Fokker Planck equation, Methods of solution and application 2nd Ed, Springer Verlag, Berlin, Heidelberg, 1989.

[44] T. D. Frank, Nonlinear Fokker-Planck equations: fundamentals and applications, Springer Science & Business Media, 2005.

[45] C. Honisch, R. Friedrich, Estimation of Kramers-Moyal coefficients at low sampling rates, Physical Review E 83 (6) (2011) 066701.

[46] M. R. R. Tabar, The Langevin Equation and Wiener Process, Springer International Publishing, Cham, 2019, pp. 39–48. doi:10.1007/978-3-030-18472-8_5.
URL https://doi.org/10.1007/978-3-030-18472-8_5

[47] M. Siefert, A. Kittel, R. Friedrich, J. Peinke, On a quantitative method to analyze dynamical and measurement noise, EPL (Europhysics Letters) 61 (4) (2003) 466.

[48] J. Gottschall, J. Peinke, On the definition and handling of different drift and diffusion estimates, New Journal of Physics 10 (8) (2008) 083034.

[49] R. Friedrich, S. Siegert, J. Peinke, M. Siefert, M. Lindemann, J. Raethjen, G. Deuschl, G. Pfister, et al., Extracting model equations from experimental data, Physics Letters A 271 (3) (2000) 217–222.

[50] D. Kleinhans, R. Friedrich, A. Nawroth, J. Peinke, An iterative procedure for the estimation of drift and diffusion coefficients of Langevin processes, Physics Letters A 346 (1-3) (2005) 42–46.

[51] J. Gradišek, I. Grabec, S. Siegert, R. Friedrich, Qualitative and quantitative analysis of stochastic processes based on measured data, I: Theory and applications to synthetic data, Journal of sound and vibration 252 (3) (2002) 545–562.

[52] J. Gradišek, E. Govekar, I. Grabec, Qualitative and quantitative analysis of stochastic processes based on measured data, II: Applications to experimental data, Journal of sound and vibration 252 (3) (2002) 563–572.

[53] P. Sura, S. T. Gille, Interpreting wind-driven Southern Ocean variability in a stochastic framework, Journal of marine research 61 (3) (2003) 313–334.

[54] S. T. Gille, Statistical characterization of zonal and meridional ocean wind stress, Journal of Atmospheric and Oceanic Technology 22 (9) (2005) 1353–1372.

[55] L. Giuggioli, T. J. McKetterick, V. Kenkre, M. Chase, Fokker–Planck description for a linear delayed Langevin equation with additive Gaussian noise, Journal of Physics A: Mathematical and Theoretical 49 (38) (2016) 384002.

[56] L. Giuggioli, Z. Neu, Fokker–Planck representations of non-Markov Langevin equations: application to delayed systems, Philosophical Transactions of the Royal Society A 377 (2153) (2019) 20180131.

[57] G. Pesce, A. McDaniel, S. Hottovy, J. Wehr, G. Volpe, Stratonovich-to-Itô transition in noisy systems with multiplicative feedback, Nature communications 4 (1) (2013) 1–7.

[58] W.-T. Lin, C.-L. Ho, Similarity solutions of the Fokker–Planck equation with time-dependent coefficients, Annals of Physics 327 (2) (2012) 386–397.

[59] L. Pichler, A. Masud, L. A. Bergman, Numerical solution of the Fokker–Planck equation by finite difference and finite element methodsa comparative study, in: Computational Methods in Stochastic Dynamics, Springer, 2013, pp. 69–85.

[60] P. Kumar, S. Narayanan, Solution of Fokker-Planck equation by finite element and finite difference methods for nonlinear systems, Sadhana 31 (4) (2006) 445–461.

[61] G. Casella, R. L. Berger, Statistical inference, Vol. 2, Duxbury Pacific Grove, CA, 2002.

[62] S. Siegert, R. Friedrich, J. Peinke, Analysis of data sets of stochastic systems, arXiv preprint cond-mat/9803250.

[63] P. Sura, J. Barsugli, A note on estimating drift and diffusion parameters from timeseries, Physics Letters A 305 (5) (2002) 304–311.

[64] M. Ragwitz, H. Kantz, Indispensable finite time corrections for Fokker-Planck equations from time series data, Physical Review Letters 87 (25) (2001) 254501.

[65] R. Friedrich, C. Renner, M. Siefert, J. Peinke, Comment on "Indispensable finite time corrections for Fokker-Planck equations from time series data", Physical Review Letters 89 (14) (2002) 149401.

[66] J. J. Stickel, Data smoothing and numerical differentiation by a regularization method, Computers & chemical engineering 34 (4) (2010) 467–475.

[67] H. Hassan, A. Mohamad, G. Atteia, An algorithm for the finite difference approximation of derivatives with arbitrary degree and order of accuracy, Journal of Computational and Applied Mathematics 236 (10) (2012) 2622–2631.

[68] S. Wojtkiewicz, L. Bergman, Numerical solution of high dimensional Fokker-Planck equations, in: 8th ASCE Specialty Conference on Probablistic Mechanics and Structural Reliability, Notre Dame, IN, USA, 2000.

[69] E. R. Homer, V. Tikare, E. A. Holm, Hybrid Potts-phase field model for coupled microstructural–compositional evolution, Computational Materials Science 69 (2013) 414–423.

[70] S. Plimpton, A. Thompson, A. Slepoy, SPPARKS kinetic monte carlo simulator (2012).

[71] A. W. Bowman, A. Azzalini, Applied smoothing techniques for data analysis: the kernel approach with S-Plus illustrations, Vol. 18, OUP Oxford, 1997.

[72] D. Gaston, C. Newman, G. Hansen, D. Lebrun-Grandie, MOOSE: A parallel computational framework for coupled systems of nonlinear equations, Nuclear Engineering and Design 239 (10) (2009) 1768–1778.

[73] A. Tran, M. Scott, J. M. Furlan, K. V. Pagalthivarthi, R. J. Visintainer, T. Wildey, aphBO-2GP-3B: A budgeted asynchronously-parallel multi-acquisition for known/unknown constrained Bayesian optimization on high-performing computing architecture, Reliability Engineering and System Safety.

[74] A. Tran, T. Wildey, S. McCann, sMF-BO-2CoGP: A sequential multi-fidelity constrained Bayesian optimization for design applications, Journal of Computing and Information Science in Engineering.

[75] A. Tran, J. Sun, J. M. Furlan, K. V. Pagalthivarthi, R. J. Visintainer, Y. Wang, 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 347 (2019) 827–852.

[76] A. Tran, M. Tran, Y. Wang, Constrained mixed-integer Gaussian mixture Bayesian optimization and its applications in designing fractal and auxetic metamaterials, Structural and Multidisciplinary Optimization (2019) 1–24.

[77] S. Plimpton, Fast parallel algorithms for short-range molecular dynamics, Journal of Computational Physics 117 (1) (1995) 1–19.

[78] A. McGaughey, M. Kaviany, Thermal conductivity decomposition and analysis using molecular dynamics simulations. part I. Lennard-Jones argon, International Journal of Heat and Mass Transfer 47 (8) (2004) 1783–1798.

[79] P. Borgelt, C. Hoheisel, G. Stell, Exact molecular dynamics and kinetic theory results for thermal transport coefficients of the Lennard-Jones argon fluid in a wide region of states, Physical Review A 42 (2) (1990) 789.

[80] A. Dawid, Z. Gburski, Interaction-induced light scattering in Lennard-Jones argon clusters: Computer simulations, Physical Review A 56 (4) (1997) 3294.

[81] K. Laasonen, S. Wonczak, R. Strey, A. Laaksonen, Molecular dynamics simulations of gas–liquid nucleation of Lennard-Jones fluid, The Journal of Chemical Physics 113 (21) (2000) 9741–9747.

[82] D. Reith, F. Müller-Plathe, On the nature of thermal diffusion in binary Lennard-Jones liquids, The Journal of Chemical Physics 112 (5) (2000) 2436–2443.

[83] M. Griebel, S. Knapek, G. Zumbusch, Numerical simulation in molecular dynamics: numerics, algorithms, parallelization, applications, Vol. 5, Springer Science & Business Media, 2007.

[84] B. Dünweg, W. Paul, Brownian dynamics simulations without Gaussian random numbers, International Journal of Modern Physics C 2 (03) (1991) 817–827.

[85] M. A. Heroux, R. A. Bartlett, V. E. Howle, R. J. Hoekstra, J. J. Hu, T. G. Kolda, R. B. Lehoucq, K. R. Long, R. P. Pawlowski, E. T. Phipps, et al., An overview of the Trilinos project, ACM Transactions on Mathematical Software (TOMS) 31 (3) (2005) 397–423.

[86] M. A. Heroux, J. M. Willenbring, A new overview of the Trilinos project, Scientific Programming 20 (2) (2012) 83–88.

[87] S. Balay, W. D. Gropp, L. C. McInnes, B. F. Smith, Efficient management of parallelism in object-oriented numerical software libraries, in: Modern software tools for scientific computing, Springer, 1997, pp. 163–202.

[88] S. Abhyankar, J. Brown, E. M. Constantinescu, D. Ghosh, B. F. Smith, H. Zhang, Petsc/ts: A modern scalable ode/dae solver library, arXiv preprint arXiv:1806.01437.

[89] S. Torquato, Statistical description of microstructures, Annual review of materials research 32 (1) (2002) 77–111.

[90] M. Groeber, S. Ghosh, M. D. Uchic, D. M. Dimiduk, A framework for automated analysis and simulation of 3D polycrystalline microstructures. Part 1: Statistical characterization, Acta Materialia 56 (6) (2008) 1257–1273.

[91] M. Groeber, S. Ghosh, M. D. Uchic, D. M. Dimiduk, A framework for automated analysis and simulation of 3D polycrystalline microstructures. Part 2: Synthetic structure generation, Acta Materialia 56 (6) (2008) 1274–1287.

[92] R. Bostanabad, Y. Zhang, X. Li, T. Kearney, L. C. Brinson, D. W. Apley, W. K. Liu, W. Chen, Computational microstructure characterization and reconstruction: Review of the state-of-the-art techniques, Progress in Materials Science 95 (2018) 1–41.

[93] Y. Liu, M. S. Greene, W. Chen, D. A. Dikin, W. K. Liu, Computational microstructure characterization and reconstruction for stochastic multiscale material design, Computer-Aided Design 45 (1) (2013) 65–76.