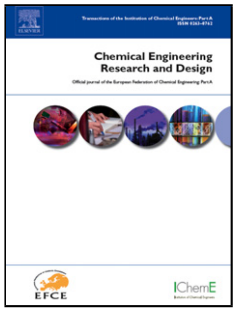
Accepted Manuscript



Title: Multilevel Monte Carlo for Noise Estimation in Stochastic Multiscale Systems

Authors: Grigoriy Kimaev, Luis A. Ricardez-Sandoval

PII: S0263-8762(18)30528-8
DOI: https://doi.org/10.1016/j.cherd.2018.10.006
Reference: CHERD 3376

To appear in:

Received date: 6-8-2018
Revised date: 24-9-2018
Accepted date: 8-10-2018

Please cite this article as: Kimaev, Grigoriy, Ricardez-Sandoval, Luis A., Multilevel Monte Carlo for Noise Estimation in Stochastic Multiscale Systems.Chemical Engineering Research and Design https://doi.org/10.1016/j.cherd.2018.10.006

Multilevel Monte Carlo for Noise Estimation in Stochastic Multiscale Systems

Grigoriy Kimaev, *Luis A. Ricardez-Sandoval

Dept. of Chemical Engineering, University of Waterloo, Waterloo, Ontario, Canada, N2L 3G1

*Corresponding author: laricard@uwaterloo.ca; Phone: +1(519)-888-4567 x38667; Fax: +1(519)-888-4347

Highlights

- Multilevel Monte Carlo sampling (MLMC) provided conservative noise estimates
- MLMC prediction accuracy could be improved with *a priori* knowledge of the system
- MLMC computational cost was an order of magnitude less than Monte Carlo sampling
- A lower bound on the achievable MLMC mean square error tolerance was observed
- Using the mean kMC timestep in the multiscale model did not affect its predictions

Abstract

The purpose of this study is to adapt Multilevel Monte Carlo (MLMC) sampling technique for random noise estimation in stochastic multiscale systems and evaluate the performance of this method. The system under consideration was a simulation of thin film formation by chemical vapour deposition, where a kinetic Monte Carlo solid-on-solid model was coupled with partial differential equations that represented mass, energy and momentum transport. The noise in the expected value of the system's observable (film roughness) was estimated using MLMC and standard Monte Carlo (MC) sampling. The MLMC technique achieved conservative estimates of noise in the observable at an order of magnitude lower computational cost than standard MC sampling. This study highlights the nuances of adapting the MLMC technique to

the stochastic multiscale system and provides insight on the benefits and challenges of using MLMC for noise estimation in stochastic multiscale systems.

Keywords: noise estimation, multilevel Monte Carlo, kinetic Monte Carlo, stochastic multiscale

1. Introduction

Many chemical engineering processes involve phenomena that occur at multiple temporal and spatial scales, for example: protein crystallization for pharmaceuticals, heterogeneous catalysis that employs supported catalysts, semiconductor doping, as well as the manufacturing of thin films for microprocessors, micro-electro-mechanical systems (MEMS), solar cells and biomedical device coatings (Christofides et al. 2009; Christofides and Armaou 2006; Crose et al. 2015; Kwon et al. 2013; Kwon, Nayhouse, Christofides, et al. 2014; Kwon, Nayhouse, Orkoulas, et al. 2014; Raimondeau and Vlachos 2000; Salciccioli et al. 2011). Some of the techniques employed in the aforementioned applications are continuous and plug flow crystallization, ion deposition by sputtering, chemical and physical vapour deposition (CVD and PVD, respectively), and plasma-enhanced chemical vapour deposition (PECVD) (Christofides and Armaou 2006; Crose et al. 2015, 2017; Kwon, Nayhouse, Christofides, et al. 2014; Kwon, Nayhouse, Orkoulas, et al. 2014; Li, Croiset, and Ricardez-Sandoval 2015).

During these deposition processes, the fluid particles are transferred by diffusion and convection over macroscale distances from the gas inlet to the surface of the wafer. Once deposition onto the substrate occurs, the mean free path of the adsorbed particles becomes comparable to the range of interactions between neighbouring particles (on the order of angstroms and nanometres), as well as the characteristic dimensions of the thin film. In addition, the rates of surface events (e.g. migration and/or catalyzed

chemical reactions) are orders of magnitude faster than the mass transfer to the substrate. Nevertheless, macroscale quantities such as inlet gas mixture concentration and substrate temperature strongly affect the microscale properties of the products (e.g. roughness, composition and thickness). Consequently, in the industrial mass production and laboratory settings, the macroscale variables are used for process optimization and control (Huang et al. 2011).

In order to better control the properties of the manufactured products and ensure compliance with industrial tolerances, various multiscale simulation algorithms have been developed to capture the inherent multiscale nature of the underlying processes (Crose et al. 2018; Kwon, Nayhouse, and Christofides 2015; Lam and Vlachos 2001; Raimondeau and Vlachos 2000; Salciccioli et al. 2011; Vlachos 1997, 2005). These algorithms rely on coupled continuum and discrete simulations where the feedbacks between the scales are explicitly accounted for; for example, the number of fluid particles that adsorb to, react on and desorb from a catalytic surface can be tracked in a simulation and used to update the boundary conditions for the equations at various length and time scales (Braatz et al. 2004, 2006; Li, Croiset, and Ricardez-Sandoval 2015; Ricardez-Sandoval 2011). The macroscale (continuum) transport phenomena are adequately captured using deterministic partial differential equations (PDEs) (Bird, Stewart, and Lightfoot 2002). However, the microscale is represented by discrete stochastic simulations such as molecular dynamics (MD), Metropolis Monte Carlo (MMC) and kinetic Monte Carlo (kMC) (Christofides et al. 2009; Frenkel and Smit 2001). To date, the MMC and kMC techniques have been more popular in multiscale algorithms because the computational intensity of MD is higher.

The need for the non-closed-form (i.e. stochastic) simulations arises because at the microscale the mean free path of the particles becomes significant (i.e. the continuum hypothesis no longer applies) and the innate discrete nature of matter must be considered: particle behaviour at the microscale is governed by the random Brownian motion and many-body effects can occur. Furthermore, microscale product properties (e.g. the local composition and uniformity of a coating) directly affect product performance;

such properties are inaccessible to macroscale calculations. Although microscale simulations are highly accurate, they become computationally intractable for experimentally-relevant scales because the domain sizes encountered in nature and in man-made devices are on the order of Avogadro's constant. Hence, the multiscale simulation algorithms aim to leverage both the computational efficiency of continuum macroscale simulations and the accuracy of the microscale simulations by "informing" the continuum scale with the results from the microscale (and vice versa).

Since microscale simulations incorporate non-closed-form expressions, noise is inherent at the discrete scale. This stochasticity also propagates into the calculations done at the macroscale and introduces variability in the simulation predictions (hence the name "*stochastic* multiscale systems"). Accounting for the noise expected in the observables would establish a greater degree of trust in the simulations' predictions, as well as minimize model-plant mismatch and enable the design of robust processes. Furthermore, since all experimentally measured data contains noise, accounting for the variability in the simulated observables would improve the simulations' ability to predict real-world scenarios. The simplest way to estimate this variability is through Monte Carlo (MC) importance sampling. It is also the most resource-intensive, because it requires prolonged sampling of the finely discretized system. The high computational cost limits the applications of MC sampling for time-sensitive applications, such as online process control. More efficient sampling and uncertainty quantification methods have been developed, such as Latin Hypercube sampling, Power Series and Polynomial Chaos Expansions (Eldred 2009). More recently, Multi Level Monte Carlo (MLMC) sampling has been developed in order to accelerate the convergence of estimates of the expected value of interest (Giles 2008).

The computational intensity of sampling a simulation system can be lessened by MLMC if the finer discretization (also referred to as "higher discretization level") causes an increase in both the precision of the results and the computational expense of obtaining the samples. This is because MLMC sampling uses many coarse samples obtained at a low computational cost to obtain a control estimate of the expected

value of the observable and then refines that approximation using progressively fewer samples of higher precision obtained at a growing computational cost. This approach causes the overall computational expenditure to be lower than extended sampling at a fine discretization, yet it was demonstrated that the results are expected to have the same level of accuracy (Giles 2008, 2015).

Note that MLMC is a sampling technique rather than a model identification method. However, if the conditions for the applicability of MLMC are met (discussed further in section 2.2), MLMC can be applied to data-driven models (Chaffart and Ricardez-Sandoval 2018; Hasenauer et al. 2015; Oladyshkin and Nowak 2012) and the reduced order models developed using the subspace identification techniques (Garg et al. 2017; Garg and Mhaskar 2017; Meidanshahi et al. 2017) for the efficient estimation of the models' observables in the presence of stochastic noise and/or parametric uncertainty.

To date, most MLMC implementations have been performed for systems comprised of closed-form continuous equations (e.g. stiff stochastic differential equations) subject to various kinds of uncertainty (Abdulle and Blumenthal 2013; Giles 2015; Mishra, Schwab, and Šukys 2012; Sauer 2013). To the extent of our knowledge, applications of MLMC sampling to stochastic multiscale systems and in the field of chemical engineering are limited. Current applications include the modelling of flows in highly heterogeneous porous media, estimation of cumulative distribution functions in stochastic oil reservoir simulations, and uncertainty quantification in surfactant/polymer enhanced-oil-recovery (Alkhatib and Babaei 2016; Guha and Tan 2017; Lu et al. 2016).

Previously, we used MLMC sampling for uncertainty quantification in three standard chemical engineering systems, namely, a mixing tank, a wastewater treatment plant, and a ternary distillation column (Kimaev and Ricardez-Sandoval 2018). We found that MLMC outperformed MC sampling, polynomial chaos and power series expansion techniques in terms of accuracy, because MLMC estimates of the observables were within the specified error tolerances from the benchmark values, unlike the results of the other

techniques. MLMC was also more computationally efficient than MC sampling. In this work, we adapt MLMC sampling to evaluate the technique's performance when it estimates the random noise in the observables of stochastic multiscale systems. To demonstrate the approach, we applied MLMC to a simulation of thin film formation by chemical vapour deposition and estimated the noise in a key observable of this system (i.e. thin film roughness at the end of the deposition process). To the extent of our knowledge, this is the first application of MLMC to a multiscale system that incorporates non-closed-form equations. In our chosen system, the discrete scale was simulated using the kMC technique. In our application of MLMC, spatial discretization is controlled by the MLMC algorithm, whereas most MLMC implementations focus on controlling the temporal discretization. We observed that using the average value of the kMC timestep produced the same thin film roughness values as the original multiscale model, especially for larger thin films. To our knowledge, this equivalence has not been reported in the literature.

The remainder of this paper is organized as follows. The next section describes the multiscale system that was used in the present analysis for the evaluation of MLMC performance. After the description of the model, an overview of the MLMC sampling technique is presented, followed by an explanation of the changes made to the multiscale model to enable the application of MLMC. Next, in the results section, we compare the amounts of time needed to approximate the maximum variability in the observable using MLMC and the standard MC sampling, as well as the accuracy of the estimates. We conclude with a discussion of the advantages and potential drawbacks of applying MLMC sampling to multiscale stochastic systems.

2. Theory and calculation

To demonstrate the principles and approach considered in this work, a case study of a stochastic multiscale system involving thin film formation by chemical vapour deposition was considered (Vlachos 1997, 2005). The model has been developed previously (Christofides and Armaou 2006; Lam and Vlachos

2001; Rasoulian and Ricardez-Sandoval 2014; Vlachos 1997), and in this work it has been used to represent an industrially-relevant stochastic multiscale chemical engineering process. An overview of the model is presented below; further details can be found elsewhere (Christofides and Armaou 2006; Kimaev and Ricardez-Sandoval 2017; Rasoulian and Ricardez-Sandoval 2014; Vlachos 1997).

2.1 Description of the stochastic multiscale model

The multiscale model consists of the kinetic Monte Carlo (kMC) simulation model coupled with partial differential equations (PDEs). The kMC model represents the microscale thin film formation by chemical vapour deposition (CVD), while the PDEs describe the macroscale mass transfer of gas particles to the substrate in the gas boundary layer above the film (Christofides and Armaou 2006; Rasoulian and Ricardez-Sandoval 2014, 2015b, 2015a, 2016; Vlachos 1997). The PDEs are necessary because the concentration, temperature and mole fraction profile of the boundary layer differ from the bulk gas phase since the gas particles continue to adsorb onto and desorb from the thin film throughout the simulation. An overview of the macroscale and fine-scale models is presented next.

2.1.1 Continuum (macroscale) model

The following momentum, energy and mass transfer equations, respectively, are used to simulate the gas boundary layer (Lam and Vlachos 2001):

$$\frac{\partial}{\partial\tau}\left(\frac{\partial f}{\partial\eta}\right)=\frac{\partial^3 f}{\partial\eta^3}+f\frac{\partial^2 f}{\partial\eta^2}+\frac{1}{2}\left[\frac{\rho_b}{\rho}-\left(\frac{\partial f}{\partial\eta}\right)^2\right] \tag{1}$$

$$\frac{\partial T}{\partial\tau}=\frac{1}{Pr}\frac{\partial^2 T}{\partial\eta^2}+f\frac{\partial T}{\partial\eta} \tag{2}$$

$$\frac{\partial x}{\partial\tau}=\frac{1}{Sc}\frac{\partial^2 x}{\partial\eta^2}+f\frac{\partial x}{\partial\eta} \tag{3}$$

The parameters used in equations (1)-(3) and other equations in the multiscale model have been described in Table 1, and their values can be found in (Kimaev and Ricardez-Sandoval 2017). The solution of equations (1)-(3) is simplified using the axisymmetric flow assumption. It is also assumed that the gas

concentration along the radial direction is uniform (Rasoulian and Ricardez-Sandoval 2014). Furthermore, the steady state assumption allows the equations of momentum and energy transfer, (1) and (2), respectively, to be solved once at the beginning of the simulation and removes the need to regularly recalculate their solutions. Only the mass transfer equation (3) is solved repeatedly throughout the simulation whenever the kMC and PDE parts of the model are coupled at the surface of the thin film through equations (10), (11) and (15).

Equations (1)-(3) are subject to the following boundary conditions in the bulk ($\eta \to \infty$):

$$T = T_{bulk} \tag{4}$$

$$\frac{\partial f}{\partial \eta} = 1 \tag{5}$$

$$x = X \tag{6}$$

and at the thin film surface ($\eta \to 0$):

$$T = T_{surface} \tag{7}$$

$$f = 0 \tag{8}$$

$$\frac{\partial f}{\partial \eta} = 0 \tag{9}$$

$$\frac{\partial x}{\partial \eta} = \frac{Sc(R_a - R_d)}{\sqrt{2a\mu_b\rho_b}} \tag{10}$$

The boundary condition (10) needs to be updated throughout the simulation. Since the macroscale calculations cannot correctly compute the rates of adsorption and desorption, R_a and R_d, respectively, the information about these rates needs to be estimated from the microscale kMC calculations.

2.1.2 Microscale kinetic Monte Carlo model

The kMC model is an on-lattice solid-on-solid model where a cubic lattice represents the thin film structure and periodic boundary conditions simulate the bulk behaviour. Only the first nearest neighbours of a particle are assumed to interact (the one atom underneath and up to four on the sides). Three surface

events are modelled: adsorption, desorption and migration. Their rates are calculated by equations (11), (12) and (13), respectively (Lam and Vlachos 2001):

$$W_a = \frac{S_0 P x|_{\eta \to 0}}{\sqrt{2\pi m R_g T C_{tot}}} N^2 \tag{11}$$

$$W_d = \sum_{i=1}^{5} M_i k_{d0} e^{-\frac{iE+E_d}{RT}} \tag{12}$$

$$W_m = \sum_{i=1}^{5} M_i k_{d0} e^{-\frac{iE+E_m}{RT}} \tag{13}$$

The choice of one of the three events is determined by the results of equations (11)-(13) and a uniform random number drawn from the open interval $(0,1)$ as per (Vlachos 1997). Whenever an event is performed, the kMC time is incremented according to equation (14), which relies on another uniform random number ς from the interval $(0,1)$ and the total rate of the system (Vlachos 1997):

$$\Delta t_{kMC} = -\frac{\ln \varsigma}{W_a + W_d + W_m} \tag{14}$$

The phenomena at the macro and micro scales affect each other: the mole fraction of the precursor adsorbing from the gas phase ($x|_{\eta \to 0}$), obtained by solving PDE (3) together with boundary conditions (6) and (10), affects the rate of adsorption (W_a) of the kMC equation (11). The rate W_a then influences Δt_{kMC} of equation (14). In turn, the kMC model keeps track of the number of adsorbed and desorbed atoms, N_a and N_d, and increments the kMC time by Δt_{kMC} after each event occurs. Eventually, when the kMC time reaches the coupling time interval $\Delta \mathcal{T}$, the macroscale rates of adsorption and desorption (R_a and R_d, respectively), are updated:

$$R_a - R_d = \frac{N_a - N_d}{2aN^2 \Delta \mathcal{T}} \tag{15}$$

The duration of the coupling time interval $\Delta \mathcal{T}$ was chosen such that it would allow a sufficient number of kMC events to occur in between the coupling instances to enable the gas concentration above the thin film surface to reach steady state (Vlachos 1997). The numeric value of $\Delta \mathcal{T}$ can be found elsewhere

(Kimaev and Ricardez-Sandoval 2017). The outcome of equation (15) is subsequently used to update the macroscale boundary condition (10), thus completing the feedback loop (i.e. the coupling) between the two scales in this model, which is summarized in Figure 1. Additional details of the implementation can be found in previous studies (Christofides and Armaou 2006; Rasoulian and Ricardez-Sandoval 2014; Vlachos 1997).

For the present analysis, the roughness R of the thin film is considered as the main observable of interest because it is susceptible to noise and it must be below a threshold for the thin film's electrical conductivity to be acceptable (Chen et al. 2017; Ketenoğlu and Ünal 2013; G. Palasantzas and Barnaś 1997; George Palasantzas 1998). Roughness can be calculated as the number of broken bonds on the surface of the film (Christofides et al. 2009; Raimondeau and Vlachos 2000):

$$R = 1 + \frac{\sum_{q=1}^{N}\sum_{z=1}^{N}\left(\left|h_{q+1,z} - h_{q,z}\right| + \left|h_{q-1,z} - h_{q,z}\right| + \left|h_{q,z+1} - h_{q,z}\right| + \left|h_{q,z-1} - h_{q,z}\right|\right)}{2N^2} \quad (16)$$

where $h_{q,z}$ is the number of gas particles adsorbed at the lattice site (q, z). Note that equation (16) relies on the absolute values of the differences in the numbers of atoms at the adjacent lattice sites.

2.2 Multilevel Monte Carlo sampling technique

The Multilevel Monte Carlo (MLMC) sampling technique aims to efficiently approximate the expected value of an observable of interest, e.g. $\mathcal{P}$, by combining the estimates obtained at different levels of accuracy and computational cost (Giles 2008). The true expected value $\mathbb{E}[\mathcal{P}]$ of a random variable $\mathcal{P}$ is given by the following equation:

$$\mathbb{E}[\mathcal{P}] = \mathbb{E}[\mathcal{P}_0] + \sum_{l=1}^{L} \mathbb{E}[\mathcal{P}_l - \mathcal{P}_{l-1}] \quad (17)$$

where L is the number of the finest discretization level, $\mathcal{P}_0$ is an estimate of $\mathcal{P}$ obtained using the coarsest discretization level ($l = 0$), $\mathcal{P}_l$ is an estimate of $\mathcal{P}$ obtained using the discretization level l, and $\mathcal{P}_l$ and $\mathcal{P}_{l-1}$ are obtained using the same Brownian path (Giles 2008). During MLMC sampling, the expected value at the coarsest level, $\mathbb{E}[\mathcal{P}_0]$, is calculated first. The $\mathbb{E}[\mathcal{P}_0]$ value is then refined with $\mathbb{E}[\mathcal{P}_l - \mathcal{P}_{l-1}]$, which are the expected values of the differences between the successive estimates of $\mathcal{P}$. It is evident that the right-hand side of equation (17) will approach $\mathbb{E}[\mathcal{P}]$ as $L \to \infty$. At the beginning of MLMC sampling, L is set to 2 to ensure that the system is sampled at least at three discretization levels ($l = 0, 1, 2$) before the convergence of the estimates of $\mathcal{P}$ is assessed, and that the variance of the estimates of $\mathcal{P}$ consistently decreases with the refinement of the discretization level. During MLMC sampling, L may need to be increased beyond 2 because samples from finer discretization levels l may be needed to decrease the variability of the estimates of $\mathcal{P}$. MLMC sampling continues until the variance of $\mathcal{P}_l - \mathcal{P}_{l-1}$ from the last three levels l satisfies ϵ, which is a user-defined upper bound of the root-mean-square error of the MLMC estimate of $\mathbb{E}[\mathcal{P}]$. Upon the completion of MLMC sampling, L becomes a finite positive integer greater than or equal to 2, such that ϵ is satisfied at levels $l = L - 2$, $L - 1$ and L.

The number of samples at every level l, i.e. $\mathcal{N}_l$, used to approximate the expected values $\mathbb{E}[\mathcal{P}_0]$ and $\mathbb{E}[\mathcal{P}_l - \mathcal{P}_{l-1}]$ is diminished by the MLMC scheme to compensate for the growing computational cost associated with higher l according to the following equation:

$$\mathcal{N}_l = \left\lceil 2\epsilon^{-2}\sqrt{\mathcal{V}_l \hbar_l}\left(\sum_{l=0}^{L}\sqrt{\frac{\mathcal{V}_l}{\hbar_l}}\right)\right\rceil \tag{18}$$

where $\mathcal{V}_l$ is the variance of the estimates $\mathcal{P}_l$, and $\hbar_l$ is the domain discretization step size at the level l. The values of $\mathcal{V}_l$ need to be known *a priori* or approximated by conducting preliminary sampling of the system of interest at several discretization levels (at least three, $l = 0, 1, 2$). It is common practice to decrease $\hbar_l$ from one discretization level to the next according to a geometric series (Giles 2008, 2015).

Thus, the decrease in the calculated $\mathcal{N}_l$ is due to the decaying $\mathcal{V}_l$ and $\hbar_l$ values. Note that to achieve smaller values of the tolerance ϵ, the $\mathcal{N}_l$ values and the overall computational cost will become larger.

It is important to mention that MLMC is only applicable to systems where increases in l lead to decreases in the variance of the estimates of $\mathcal{P}$. Furthermore, to achieve computational savings with MLMC sampling, the cost of computing a sample with a finer discretization should exceed the cost associated with coarser discretization. As per the MLMC Theorem, the key condition for the applicability of MLMC is for the rate of convergence of the estimates of $\mathcal{P}$ to be greater than or equal to half of either the rate of convergence of the variance of the estimates of $\mathcal{P}$, or the rate of increase in computational complexity due to discretization refinement, whichever of the latter two rates is smaller (Giles 2008, 2015; Kimaev and Ricardez-Sandoval 2018). In section 2.3 (below), it is demonstrated that these criteria are met by the stochastic multiscale system chosen for this study. More comprehensive discussions of the factors that determine the applicability of MLMC sampling, as well as the foundational MLMC Theorem, can be found elsewhere (Aslett, Nagapetyan, and Vollmer 2017; Giles 2008, 2015). Also, a thorough description and a tutorial example of the MLMC sampling scheme can be found in (Kimaev and Ricardez-Sandoval 2018). Although the MLMC technique offers substantial time savings while preserving the accuracy, the method is not without shortcomings. The algorithm is heuristic, and its convergence is not guaranteed. The greatest time savings can be achieved only with the optimal selection of domain discretization at the $l = 0$ level. If the samples at $l = 0$ exhibit a very high variability ($\mathcal{V}_0$) due to a very coarse discretization, excessive time will be spent at $l = 0$ because $\mathcal{N}_0$ value will be high, and the overall computational cost will increase. However, when the discretization at $l = 0$ is too fine, additional MLMC levels will have to use even finer domain discretization, which will result in higher computational costs at each level and therefore increase the overall cost of MLMC sampling. A further discussion of the potential drawbacks of MLMC can be found elsewhere (Giles 2008, 2015; Kimaev and Ricardez-Sandoval 2018).

2.3 MLMC sampling applied to stochastic multiscale systems

When applying the MLMC sampling technique to a system, it is necessary to identify the key discretization domains that affect the variability of the observable values generated by the system, as well as the computational cost associated with the level of accuracy. The majority of current MLMC implementations discretize the time domains of their respective simulation systems because the systems are represented by time-dependent differential equations and smaller integration timesteps yield results of lower variability. In the case of stochastic multiscale systems, domains other than temporal (e.g. spatial) may have a greater influence on the accuracy of simulations' outcomes. For example, in systems that combine discrete particle simulations (e.g. kMC) with continuous differential equations, the number of particles considered in the discrete simulation domain may have a stronger impact on the noise in the observables than the integration timestep used to integrate the continuous differential equations.

Once the discretization domain is chosen, the sequence for discretization level refinement needs to be selected. The levels should be chosen such that the variability of the observables is reduced in an optimal way. On the one hand, a slow refinement of the discretization would reduce the variance, but the overall computational cost would be unnecessarily high since too many levels would be sampled by the MLMC scheme. On the other hand, a drastic refinement of discretization could quickly make higher discretization levels (and the total cost) computationally intractable; for example, decreasing the integration timestep of a large system of differential equations by an order of magnitude with every additional level could make the costs of sampling prohibitive after a few levels have been added by the MLMC scheme. For efficient MLMC sampling, a geometric sequence is commonly used for discretization level refinement (Giles 2015). The optimal value of the refinement factor is system-specific (Giles 2008, 2015):

$$\mathcal{M} = \frac{\Delta t_{l+1}}{\Delta t_l} \tag{19}$$

where $\mathcal{M}$ is the optimal refinement factor (e.g. ½, ¼), and Δt_l and Δt_{l+1} represent the integration timesteps used by MLMC while sampling a time-dependent system at the successive discretization levels l and $l+1$, respectively.

In the stochastic multiscale kMC-PDE model considered in this work, the size of the lattice of the kMC thin film model (represented by N, see section 2.1.2) directly affects the variance of the roughness observable at the end of the batch, as demonstrated in Figure 2. Using a relatively large lattice (e.g. $N = 100$) substantially decreases the noise in the observable (i.e. surface roughness); however, the simulation time grows with lattice size. For example, on an Intel® Core™ i7-4770 CPU @ 3.40 GHz with 32 GB of RAM, a multiscale simulation from $0\ s$ to $100\ s$ at $T = 800\ K$ would take approximately 18.2 seconds of CPU time for $N = 30$, 122.7 seconds for $N = 100$, and 472.5 seconds for $N = 150$. The relationships between lattice size, observable's variability and computational cost indicate that MLMC sampling is indeed applicable to this stochastic multiscale system.

The current heuristic approach to approximating roughness is to use six $N = 30$ lattices, find the roughness time trajectories from each lattice, find the average trajectory, and use it as an adequate representation of results that would be obtained from $N = 100$ and $N = 150$ simulations, as illustrated in Figure 3 (Christofides et al. 2009). Substantial CPU time savings can be achieved when the simulations are run in parallel. MLMC sampling can also take advantage of parallel computing, but one can approach the selection of the number and size of lattices in a more objective manner because MLMC tracks the decay in the variability of the estimates of the observable with each discretization level and adjusts the number of samples accordingly, as per equations (17) and (18).

Since larger thin film lattice sizes impact both the observable's estimates and the computational cost, the N values can be considered as discretization levels in the spatial domain. To adapt the MLMC sampling algorithm to the present multiscale model, we established an empirical relationship using least squares

regression between the lattice size N, temperature T, and the kMC timestep Δt_{kMC}, since they are already related through the kMC rate of adsorption W_a. This rate is calculated based on $\sqrt{T}$ and N^2 in equation (11) and is inversely related to Δt_{kMC} through equation (14). In this work, only N and T were manipulated. Thus, an empirical relationship was established using N and T as the independent variables. Note that since ς of equation (14) is a uniform random number, variability is present in the kMC timestep. Hence, N and T are related to the mean value of the kMC timestep ($\widehat{\Delta t}_{kMC}$) as follows:

$$\widehat{\Delta t}_{kMC} = (\alpha T + \beta) N^{-2} \tag{20}$$

where the α and β model coefficients were found by least squares regression. This model was validated by examining its quality of fit to the mean kMC timestep values collected from the multiscale model at various N and T values (discussed in detail in section 3.1 below). For the purposes of this study, equation (14) of the kMC PDE model was replaced with equation (20). Note that if another stochastic multiscale model were considered for MLMC sampling, an empirical relationship different from equation (20) may need to be established.

Next, it was necessary to establish the rule for selecting the kMC lattice sizes to enable MLMC sampling to systematically define each discretization level. Substituting equation (20) into equation (19) yields:

$$\mathcal{M} = \frac{(\alpha T + \beta) N_{l+1}^{-2}}{(\alpha T + \beta) N_l^{-2}} \tag{21}$$

which, upon simplification, results in:

$$\frac{N_{l+1}}{N_l} = \sqrt{\mathcal{M}^{-1}} \tag{22}$$

Equation (22) can be further rearranged to yield the rule for selecting the lattice sizes N_l during MLMC sampling of the modified kMC PDE model:

$$N_l = \left\lceil N_0 \times \left(\sqrt{\mathcal{M}^{-1}}\right)^l \right\rceil \tag{23}$$

where N_0 is the user-defined minimum lattice size used by MLMC sampling. Equation (23) ensures a systematic selection of discretization levels of the spatial domain of the kMC simulation: the lattice size at an MLMC discretization level $l > 0$ will correspond to the $\widehat{\Delta t}_{kMC}$ value that is an $\mathcal{M}$ factor refinement of the mean kMC timestep at the preceding discretization level, in harmony with the existing MLMC implementations (Giles 2008, 2015; Giles, Nagapetyan, and Ritter 2015; Higham 2015).

3. Results and discussion

3.1 Establishing the α and β constants

To estimate the α and β coefficients of equation (20), Δt_{kMC} values of equation (14) were collected from various simulations of the original, unmodified kMC PDE model. In total, 105 simulations were conducted, where the lattice sizes ranged from 20 to 120, inclusive, with increments of 5, at temperatures from $800\ K$ to $1200\ K$, inclusive, with increments of $100\ K$. The temperature range was selected based on the expected operating conditions for an actual thin film deposition system. The mean kMC timestep values, $\widehat{\Delta t}_{kMC}$, were calculated for each of the 105 simulations and the results have been plotted in Figure 4 (red markers). It can be seen that all 105 $\widehat{\Delta t}_{kMC}$ values were distinct from each other. Their uniqueness was further verified from their statistical distributions. Since each simulation was run from $0\ s$ to $100\ s$ with the coupling interval $\Delta\mathcal{T} = 0.1\ s$, the kMC and PDE scales were coupled 1000 times in every simulation. The mean kMC timestep values were calculated for every coupling interval of every simulation, and their histograms were plotted (one histogram per simulation). The statistical distributions from all simulations were normal or close to normal, with sharp peaks that coincided with the data (red dots) presented in Figure 4 and very short tails that did not overlap with distribution tails from other simulations (the distributions were not shown for brevity). Thus, it was concluded that the variability of the mean kMC timesteps was sufficiently low to be neglected.

Next, the α and β coefficient values of equation (20) were obtained from least squares regression ($-7.594 \times 10^{-4}\ s \cdot K^{-1}$ and $1.106\ s$, respectively). The resulting surface plot, where $\widehat{\Delta t}_{kMC}$ was calculated according to equation (20), has also been presented in Figure 4. A very close agreement between the two datasets was demonstrated: the average Δt_{kMC} of the original model (red markers in Figure 4) and $\widehat{\Delta t}_{kMC}$ from equation (20) (surface plot in Figure 4) have negligible residual errors (between -4.4% and 1.7%, not shown for brevity). Therefore, the dataset considered for N and T was sufficient to obtain an accurate empirical relationship.

Next, the method of calculating the kMC timestep in the multiscale model was modified: equation (14) was replaced with equation (20). To verify that the modified kMC model can produce sufficiently accurate results, multiscale simulations were conducted with lattice sizes $N = 50$ and $N = 100$ at the temperatures of $940\ K$ and $1140\ K$. These temperature values were chosen because they were not used to identify the coefficients of equation (20) and because they represent the two regimes of the multiscale model, the lower-temperature adsorption-dominated regime that produces thin films with higher roughness ($940\ K$) and the higher-temperature migration-dominated regime that results in smoother thin films ($1140\ K$). For each temperature, two kinds of simulations were carried out: with the original multiscale model where Δt_{kMC} was calculated by equation (14) after each kMC event, and with the modified multiscale model where the fixed $\widehat{\Delta t}_{kMC}$ was calculated from equation (20) once at the beginning of the simulation and used throughout without alterations. The comparison is presented in Figure 5, with 8 roughness time trajectories from the original model and 1 representative roughness trajectory (no averaging) from the modified model in each subplot. Clearly, the replacement of equation (14) by equation (20) preserved the accuracy of predictions. The predictions from the modified model approached the average of the time trajectories obtained from the original model, especially for larger lattice sizes ($N = 100$), as shown in Figure 5. Note that the use of a fixed kMC timestep did not affect the computational costs. Additional testing has been performed to corroborate this observation (omitted for

brevity). To the extent of our knowledge, such a result has not been reported for this stochastic multiscale system. Also, note that fixing the kMC timestep did not eliminate the noise from the roughness time trajectories, since the kMC portion of the multiscale model continued to use uniform random numbers to choose among the adsorption, desorption and migration surface events, as discussed in section 2.1.2.

3.2 Estimating the noise via MC and MLMC sampling

To compare the computational cost and accuracy of estimating the noise in the roughness observable of the multiscale model by the two sampling techniques, Monte Carlo and Multilevel Monte Carlo, multiscale simulations were conducted with the original model and the modified model (with equation (20) substituted for equation (14)) at three temperature settings: $940\ K$, $1040\ K$ and $1140\ K$. The focus of this work was to estimate the maximum noise in the average of the last five roughness values (i.e. at seconds 96-100, inclusive, of each simulation). The average is denoted from henceforth by R_f.

While MC sampling relied on the original model, the MLMC scheme used the modified version, i.e. the model that uses a fixed kMC timestep. MC sampling was performed using $N = 100$. Larger lattices were not sampled because the substantial increase in computational cost was not justified by the observed changes in the roughness trajectory. For example, a negligible difference can be observed between $N = 100$ and $N = 150$ in Figure 3, yet the computational cost is more than 3 times higher for $N = 150$. Throughout MC sampling, the maximum and minimum observed R_f values, as well as the difference between them, were updated whenever a new batch of 80 multiscale simulations finished running. In total, 6000 simulations were performed at each temperature.

MLMC sampling used equation (23) as the rule for setting the spatial domain discretization at each MLMC level. The minimum lattice size used in equation (23), i.e. N_0, was set to 20, since that was the smallest N value that did not cause extreme noise in the predictions made by the present kMC PDE model. The

optimal refinement factor ($\mathcal{M}$) was set to ½ because it was the smallest value that did not cause MLMC to sample computationally intractable lattice sizes at higher discretization levels.

MLMC sampling was performed for each of the three temperatures; the root-mean-square error tolerance ϵ was set to 0.3 and 0.2 at each temperature. To ensure adequate sampling, 64 MLMC runs were performed at each combination of the temperature and ϵ values. Once all MLMC sampling runs were completed, the differences between the maximum and minimum estimated R_f values were calculated at each setting. Figure 6 and Table 2 compare the R_f noise bounds obtained by MC and MLMC sampling.

Based on the results summarized in Figure 6, it was observed that MC sampling exhibited creeping noise in roughness values due to the asymptotic behaviour of $\max R_f - \min R_f$ values. MC sampling achieved noise bounds of 0.679 for $T = 940\ K$, 0.289 for $1040\ K$ and 0.140 for $1140\ K$. It should be noted that these noise estimates could increase with additional MC sampling. In this study, the computational cost of MC became prohibitive after such a long sampling time (i.e. over 200 hours).

From Figure 6 and Table 2, it was apparent that MLMC could produce conservative estimates of noise in R_f values and offered an order of magnitude time savings compared to MC sampling: the computational requirements of MLMC were between 6.5 hours (at $1140\ K$, $\epsilon = 0.3$) and 38.2 hours (at $940\ K$, $\epsilon = 0.2$), whereas the MC sampling time ranged from 208.3 hours (at $T = 940\ K$) to 349.8 hours (at $T = 1140\ K$), as per Table 2. It could also be observed that the MLMC and MC results at $940\ K$ and $1140\ K$ showed consistent trends: all MLMC and MC noise estimates at $1140\ K$ were smaller than at $940\ K$ (e.g. for $\epsilon = 0.2$, MLMC results were 0.809 at $940\ K$ and 0.307 at $1140\ K$, while MC results were 0.679 at $940\ K$ and 0.140 at $1140\ K$). However, at $1040\ K$, the MC estimate (i.e. 0.289) was between the MC noise values from $940\ K$ and $1140\ K$ (0.679 and 0.140, respectively), while MLMC produced a result higher than at $940\ K$ (for $\epsilon = 0.3$, at $1040\ K$ the result was 1.175, higher than 1.054 at $940\ K$). Thus, at $1040\ K$, MLMC diverged from the trends in the noise bounds observed at 940 K and 1140 K. One possible explanation to

this behaviour is that the multiscale system started to transition into the lower-noise migration-dominated regime and caused MLMC R_f noise estimate to become overly conservative. This issue could be mitigated to some extent by decreasing ϵ from 0.3 to 0.2 at $T = 1040\ K$, but for $\epsilon = 0.2$ the MLMC result at $1040\ K$ was still very close to the value at $940\ K$ (0.803 versus 0.809, respectively). The obvious shortcoming of this approach to tuning the ϵ value is that to obtain a more accurate noise estimate, the selection of ϵ has to be guided by some *a priori* knowledge of the system behaviour. Thus, despite the computational time savings offered by MLMC sampling, MC sampling may still need to be conducted for a short amount of time to obtain an approximation of the noise that can validate the estimates from MLMC. The greatest computational savings from MLMC were observed in the low-noise regime, i.e. at $1140\ K$, where MLMC required at most 9.8 hours of CPU time (for $\epsilon = 0.2$), whereas MC sampling took 349.8 hours. Since the CPU times required by MLMC were on the order of hours, online applications of MLMC for stochastic systems would still be challenging. However, the usage of high performance computing clusters could enable online applications by leveraging the parallelization of MLMC sampling. Furthermore, using larger ϵ values could also reduce the computational costs.

Note that smaller ϵ were considered (e.g. 0.1). At $T = 940\ K$ and $1040\ K$, setting ϵ to 0.1 caused MLMC to sample large, computationally intractable thin film lattices. Such MLMC sampling runs were terminated before completion and their results could not be obtained. At $1140\ K$, MLMC runs with $\epsilon = 0.1$ were able to achieve roughness estimates because the multiscale system was in the low-noise regime at that temperature, but the approximations did not improve upon the presented results. This behaviour of MLMC is different from what we observed previously for closed-form systems (Kimaev and Ricardez-Sandoval 2018). This result suggests that there is a lower bound on the ϵ error tolerance that can be achieved by MLMC for the case of stochastic systems.

Furthermore, it is imperative for MLMC to use the same Brownian path when the samples of the observable are obtained at two subsequent discretization levels, as discussed in section 2.2 and in (Giles

2008). When MLMC is applied to systems represented by continuous equations, the discretization level is the only source of discrepancies between the samples. Due to the application of MLMC sampling to a stochastic multiscale system, an additional source of variability arises: the inherent noise associated with the non-closed-form expressions will cause deviations in the results of different simulation runs even when all parameter values and the discretization level are the same. To ensure that MLMC uses the same Brownian path when sampling the observables of a stochastic multiscale system at two subsequent discretization levels, the same seed must be provided to the random number generators in the simulations performed at both discretization levels.

Note that multiple approaches for adapting the multiscale model for MLMC sampling were explored in this study. The coupling timestep $\Delta\mathcal{T}$ between the kMC and PDE scales was considered as the discretization level in MLMC sampling. However, when $\Delta\mathcal{T}$ was small, it resulted in longer simulation times and produced the same roughness trajectories. However, when $\Delta\mathcal{T}$ was large, it caused discontinuities in the roughness time trajectory, indicating numerical instability encountered at the coupling of the two scales. In addition, without a systematic way of selecting the lattice sizes, MLMC runs could become excessively long, since the simulation times of individual multiscale simulations within the MLMC scheme could require computationally prohibitive lattice sizes at higher MLMC discretization levels.

4. Conclusions

In this study, the Multilevel Monte Carlo (MLMC) sampling technique was adapted to stochastic multiscale systems. As a case study, MLMC was applied to a model that represented thin film formation by chemical vapour deposition. Unlike in most other MLMC implementations, which discretize the time domain, in this work the spatial discretization domain was used in the MLMC scheme. MLMC and Monte Carlo (MC) sampling techniques were used to estimate the noise in the system's observable, i.e. roughness of the thin film. The greatest computational efficiency of MLMC was observed when the system was in the low-

noise regime. However, while the potential computational time savings offered by MLMC were an order of magnitude compared to MC sampling, MLMC overestimated the noise values obtained from MC. The tuning of the root-mean-square-error tolerance (ϵ) of the MLMC technique improved the accuracy of the noise estimates, but it relied on *a priori* knowledge of the system's behaviour. It was also found that unlike for the case of closed-form systems of equations, in stochastic multiscale systems MLMC could not satisfy every imposed ϵ at every set of simulation conditions, which suggests the existence of a lower bound on the error tolerance.

Acknowledgements

The authors would like to gratefully acknowledge the Natural Sciences and Engineering Research Council of Canada (NSERC) for the financial support.

References

Abdulle, Assyr, and Adrian Blumenthal. 2013. "Stabilized Multilevel Monte Carlo Method for Stiff Stochastic Differential Equations." *Journal of Computational Physics* 251: 445–60.

Alkhatib, A, and M Babaei. 2016. "Applying the Multilevel Monte Carlo Method for Heterogeneity-Induced Uncertainty Quantification of Surfactant / Polymer Flooding." *SPE Journal* 21(4): 1192–1203.

Aslett, Louis J.M., Tigran Nagapetyan, and Sebastian J. Vollmer. 2017. "Multilevel Monte Carlo for Reliability Theory." *Reliability Engineering and System Safety* 165(March): 188–96.

Bird, R. Byron, Warren E. Stewart, and Edwin N. Lightfoot. 2002. *Transport Phenomena*. 2nd ed. John Wiley & Sons.

Braatz, R D et al. 2006. "Perspectives on the Design and Control of Multiscale Systems." *Journal of Process Control* 16(3): 193–204.

Braatz, R D, Richard C. Alkire, Effendi Rusli, and Timothy O. Drews. 2004. "Multiscale Systems Engineering with Applications to Chemical Reaction Processes." *Chemical Engineering Science* 59(22–23): 5623–28.

Chaffart, Donovan, and Luis A Ricardez-Sandoval. 2018. "Optimization and Control of a Thin Film Growth Process: A Hybrid First Principles/Artificial Neural Network Based Multiscale Modelling Approach." *Computers and Chemical Engineering* (In press). https://doi.org/10.1016/j.compchemeng.2018.08.029.

Chen, Zhiqiang, Wenchao Tian, Xiaotong Zhang, and Yongkun Wang. 2017. "Effect of Deposition Parameters on Surface Roughness and Consequent Electromagnetic Performance of Capacitive RF MEMS Switches: A Review." *Journal of Micromechanics and Microengineering* 27(11).

Christofides, Panagiotis D., and Antonios Armaou. 2006. "Control and Optimization of Multiscale Process Systems." *Computers and Chemical Engineering* 30(10–12): 1670–86.

Christofides, Panagiotis D., Antonios Armaou, Yiming Lou, and Amit Varshney. 2009. *Control and Optimization of Multiscale Process Systems*. 1st ed. ed. William S. Levine. New York, NY: Birkhäuser Boston.

Crose, Marquis et al. 2015. "Multiscale Modeling and Operation of PECVD of Thin Film Solar Cells." *Chemical Engineering Science* 136: 50–61.

Crose, Marquis, Joseph Sang-Il Kwon, Anh Tran, and Panagiotis D. Christofides. 2017. "Multiscale Modeling and Run-to-Run Control of PECVD of Thin Film Solar Cells." *Renewable Energy* 100: 129–40.

Crose, Marquis, Weiqi Zhang, Anh Tran, and Panagiotis D. Christofides. 2018. "Multiscale Three-Dimensional CFD Modeling for PECVD of Amorphous Silicon Thin Films." *Computers and Chemical Engineering* 113: 184–95.

Eldred, M S. 2009. "Recent Advances in Non-Intrusive Polynomial Chaos and Stochastic Collocation Methods for Uncertainty Analysis and Design." In *AIAA Structures, Structural Dynamics and Materials Conference*, Palm Springs, California, 1–37.

Frenkel, D, and B Smit. 2001. *Understanding Molecular Simulation: From Algorithms to Applications*. 2nd ed. Academic Press.

Garg, Abhinav et al. 2017. "Subspace-Based Model Identification of a Hydrogen Plant Startup Dynamics." *Computers and Chemical Engineering* 106: 183–90.

Garg, Abhinav, and Prashant Mhaskar. 2017. "Subspace Identification-Based Modeling and Control of Batch Particulate Processes." *Industrial & Engineering Chemistry Research* 56(26): 7491–7502.

Giles, Michael. 2008. "Multi-Level Monte Carlo Path Simulation." *Operations Research* 56(3): 607–17.

Giles, Michael. 2015. "Multilevel Monte Carlo Methods." *Acta Numerica* 24: 259–328.

Giles, Michael, Tigran Nagapetyan, and Klaus Ritter. 2015. "Multilevel Monte Carlo Approximation of Distribution Functions and Densities." *SIAM/ASA J. Uncertainty Quantification* 3: 267–95.

Guha, Nilabja, and Xiaosi Tan. 2017. "Multilevel Approximate Bayesian Approaches for Flows in Highly Heterogeneous Porous Media and Their Applications." *Journal of Computational and Applied Mathematics* 317: 700–717.

Hasenauer, Jan, Nick Jagiella, Sabrina Hross, and Fabian J. Theis. 2015. "Data-Driven Modelling of Biological Multi-Scale Processes." *Journal of Coupled Systems and Multiscale Dynamics* 3(2): 101–21.

Higham, Desmond J. 2015. "An Introduction to Multilevel Monte Carlo for Option Valuation." *International Journal of Computer Mathematics* 92(12): 2347–60.

Huang, Jianqiao, Xinyu Zhang, Gerassimos Orkoulas, and Panagiotis D Christofides. 2011. "Dynamics and Control of Aggregate Thin Film Surface Morphology for Improved Light Trapping: Implementation on a Large-Lattice Kinetic Monte Carlo Model." *Chemical Engineering Science* 66(23): 5955–67.

Ketenoğlu, D., and B. Ünal. 2013. "Influence of Surface Roughness on the Electrical Conductivity of Semiconducting Thin Films." *Physica A: Statistical Mechanics and Its Applications* 392(14): 3008–17.

Kimaev, Grigoriy, and Luis A Ricardez-Sandoval. 2017. "A Comparison of Efficient Uncertainty Quantification Techniques for Stochastic Multiscale Systems." *AIChE Journal* 63(8): 3361–73.

Kimaev, Grigoriy, and Luis A Ricardez-Sandoval. 2018. "Multilevel Monte Carlo Applied to Chemical Engineering Systems Subject to Uncertainty." *AIChE Journal* 64(5): 1651–61.

Kwon, Joseph Sang-Il, Michael Nayhouse, and Panagiotis D. Christofides. 2015. "Multiscale, Multidomain Modeling and Parallel Computation: Application to Crystal Shape Evolution in Crystallization." *Industrial & Engineering Chemistry Research* 54(47): 11903–11914.

Kwon, Joseph Sang-Il, Michael Nayhouse, Panagiotis D. Christofides, and Gerassimos Orkoulas. 2013. "Modeling and Control of Protein Crystal Shape and Size in Batch Crystallization." *AIChE Journal* 59(7): 2317–27.

Kwon, Joseph Sang-Il, Michael Nayhouse, Panagiotis D. Christofides, and Gerassimos Orkoulas. 2014. "Modeling and Control of Crystal Shape in Continuous Protein Crystallization." *Chemical Engineering Science* 107: 47–57.

Kwon, Joseph Sang-Il, Michael Nayhouse, Gerassimos Orkoulas, and Panagiotis D. Christofides. 2014.

"Crystal Shape and Size Control Using a Plug Flow Crystallization Configuration." *Chemical Engineering Science* 119: 30–39.

Lam, R., and D. Vlachos. 2001. "Multiscale Model for Epitaxial Growth of Films: Growth Mode Transition." *Physical Review B* 64(3): 035401.

Li, Jingde, Eric Croiset, and Luis Ricardez-Sandoval. 2015. "Carbon Nanotube Growth: First-Principles-Based Kinetic Monte Carlo Model." *Journal of Catalysis* 326: 15–25.

Lu, Dan, Guannan Zhang, Clayton Webster, and Charlotte Barbier. 2016. "An Improved Multilevel Monte Carlo Method for Estimating Probability Distribution Functions in Stochastic Oil Reservoir Simulations." *Water Resources Research* 52(12): 9642–60.

Meidanshahi, Vida, Brandon Corbett, Thomas A. Adams II, and Prashant Mhaskar. 2017. "Subspace Model Identification and Model Predictive Control Based Cost Analysis of a Semicontinuous Distillation Process." *Computers and Chemical Engineering* 103: 39–57.

Mishra, S., Ch Schwab, and J. Šukys. 2012. "Multi-Level Monte Carlo Finite Volume Methods for Nonlinear Systems of Conservation Laws in Multi-Dimensions." *Journal of Computational Physics* 231(8): 3365–88.

Oladyshkin, S., and W. Nowak. 2012. "Data-Driven Uncertainty Quantification Using the Arbitrary Polynomial Chaos Expansion." *Reliability Engineering and System Safety* 106: 179–90.

Palasantzas, G., and J. Barnaś. 1997. "Surface-Roughness Fractality Effects in Electrical Conductivity of Single Metallic and Semiconducting Films." *Physical Review B* 56(12): 7726–31.

Palasantzas, George. 1998. "Surface Roughness and Grain Boundary Scattering Effects on the Electrical Conductivity of Thin Films." *Physical Review B* 58(15): 9685–88.

Raimondeau, S, and D. G. Vlachos. 2000. "Low-Dimensional Approximations of Multiscale Epitaxial Growth Models for Microstructure Control of Materials." *Journal of Computational Physics* 160(2): 564–76.

Rasoulian, Shabnam, and Luis A. Ricardez-Sandoval. 2015a. "A Robust Nonlinear Model Predictive Controller for a Multiscale Thin Film Deposition Process." *Chemical Engineering Science* 136: 38–49.

Rasoulian, Shabnam, and Luis A. Ricardez-Sandoval. 2015b. "Robust Multivariable Estimation and Control in an Epitaxial Thin Film Growth Process under Uncertainty." *Journal of Process Control* 34: 70–81.

Rasoulian, Shabnam, and Luis A. Ricardez-Sandoval. 2016. "Stochastic Nonlinear Model Predictive Control Applied to a Thin Film Deposition Process under Uncertainty." *Chemical Engineering Science* 140: 90–103.

Rasoulian, Shabnam, and Luis Alberto Ricardez-Sandoval. 2014. "Uncertainty Analysis and Robust Optimization of Multiscale Process Systems with Application to Epitaxial Thin Film Growth." *Chemical Engineering Science* 116: 590–600.

Ricardez-Sandoval, Luis A. 2011. "Current Challenges in the Design and Control of Multiscale Systems." *The Canadian Journal of Chemical Engineering* 89(6): 1324–41.

Salciccioli, M., M. Stamatakis, S. Caratzoulas, and D. G. Vlachos. 2011. "A Review of Multiscale Modeling

of Metal-Catalyzed Reactions: Mechanism Development for Complexity and Emergent Behavior." *Chemical Engineering Science* 66(19): 4319–55.

Sauer, Timothy. 2013. "Computational Solution of Stochastic Differential Equations." *Wiley Interdisciplinary Reviews: Computational Statistics* 5(5): 362–71.

Vlachos, DG. 1997. "Multiscale Integration Hybrid Algorithms for Homogeneous–Heterogeneous Reactors." *AIChE Journal* 43(11): 3031–41.

Vlachos, DG. 2005. "A Review of Multiscale Analysis: Examples from Systems Biology, Materials Engineering, and Other Fluid–Surface Interacting Systems." *Advances in Chemical Engineering* 30: 1–61.

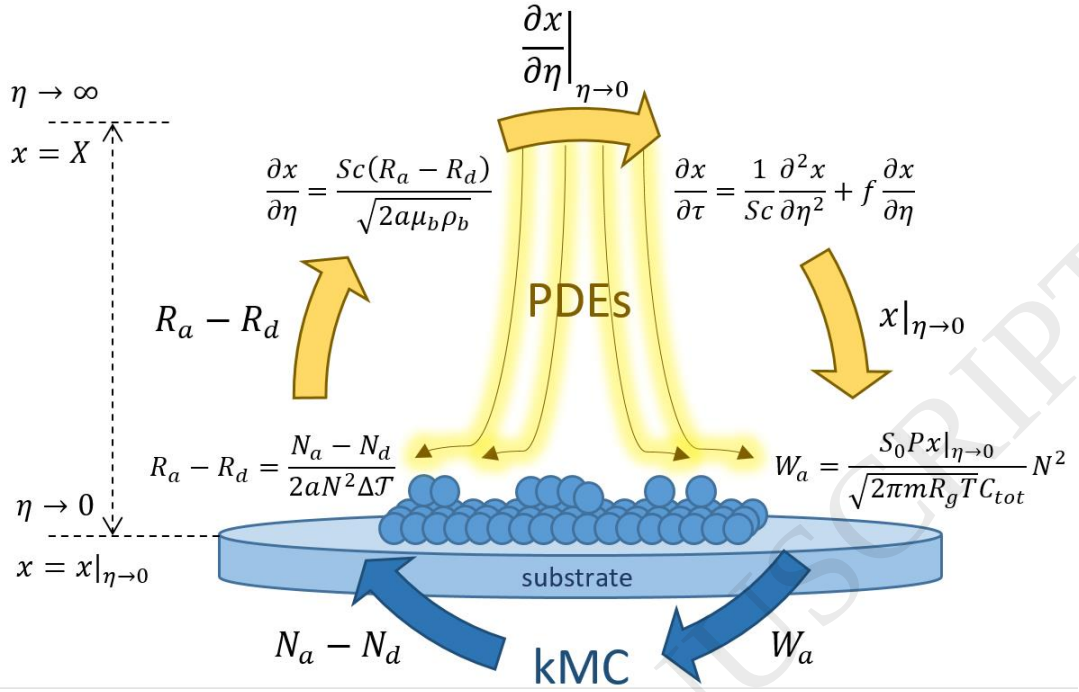


Figure 1: The multiscale feedback loop between equations (3), (11), (15) and (10) (from the top, in clockwise order).

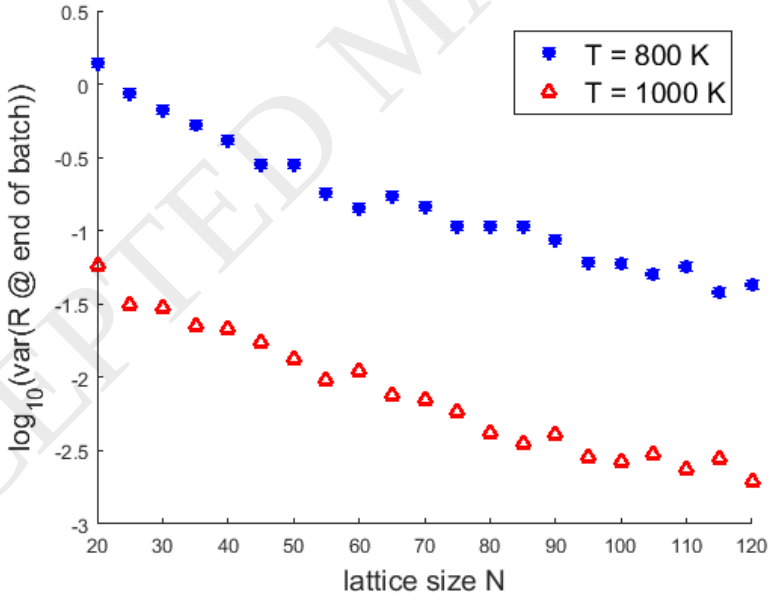


Figure 2: The decay of variance in roughness at the end of the batch due to increasing thin film lattice size. For brevity, results at only two temperature values have been shown.

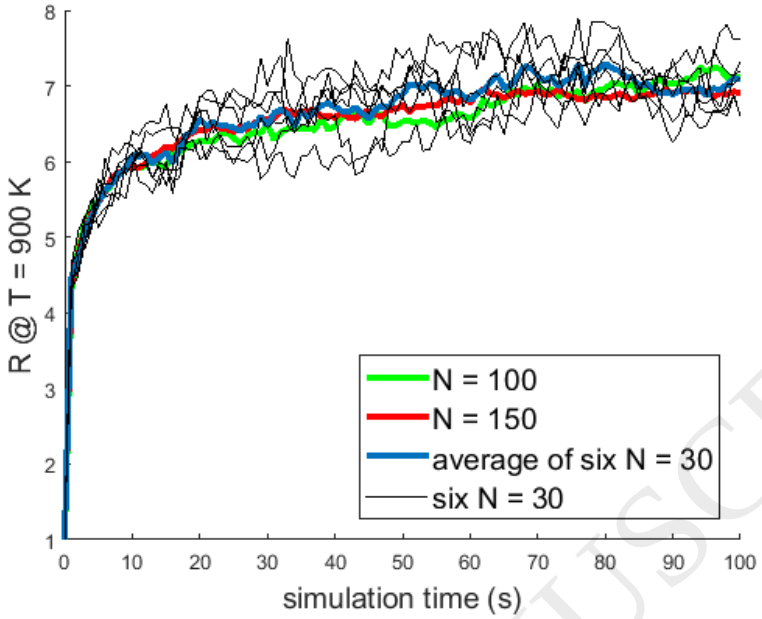


Figure 3: A comparison of roughness time trajectories from different thin film lattice sizes.

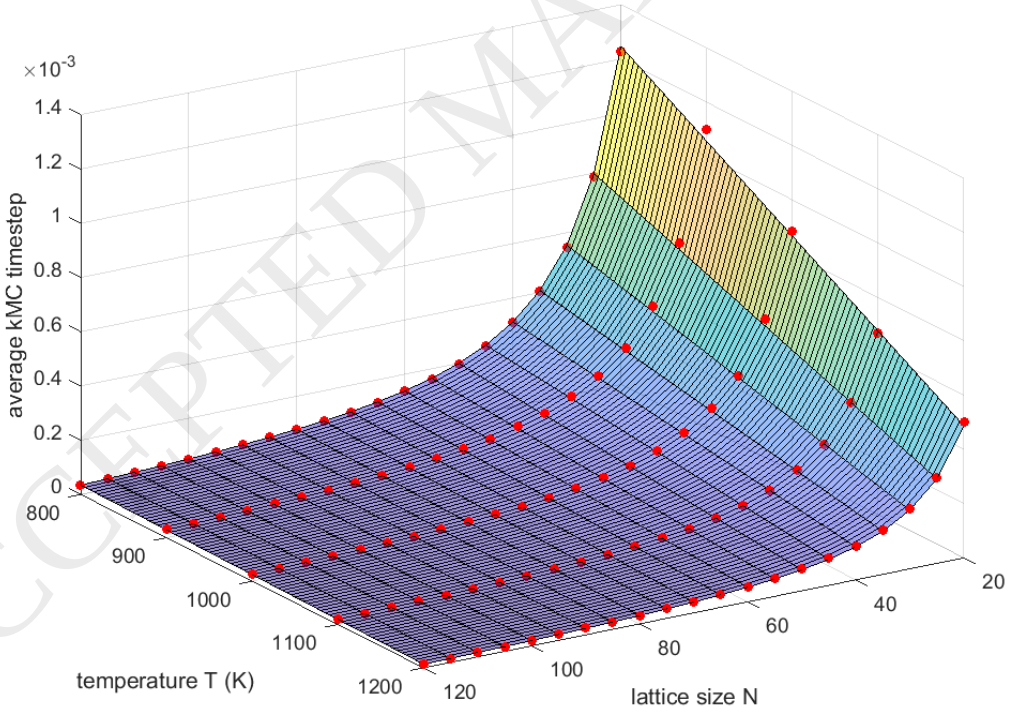


Figure 4: A comparison between the mean kMC timesteps obtained from simulations (red markers) and the kMC timesteps calculated from equation (20) (surface plot). Note that α was found to be $-7.594 \times 10^{-4}\ s \cdot K^{-1}$ and β was $1.106\ s$.

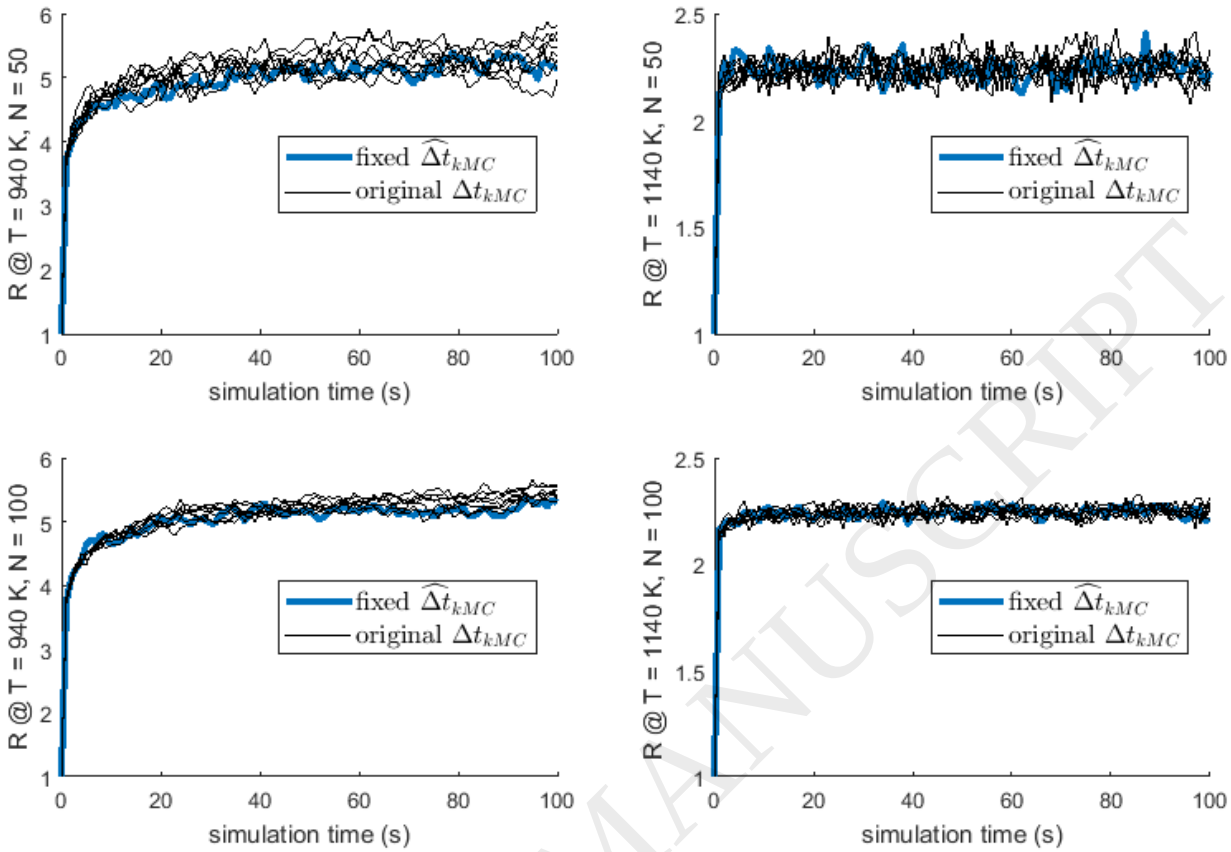


Figure 5: A comparison between the time trajectories of roughness produced by the original multiscale simulation and the modified model with a fixed kMC timestep.

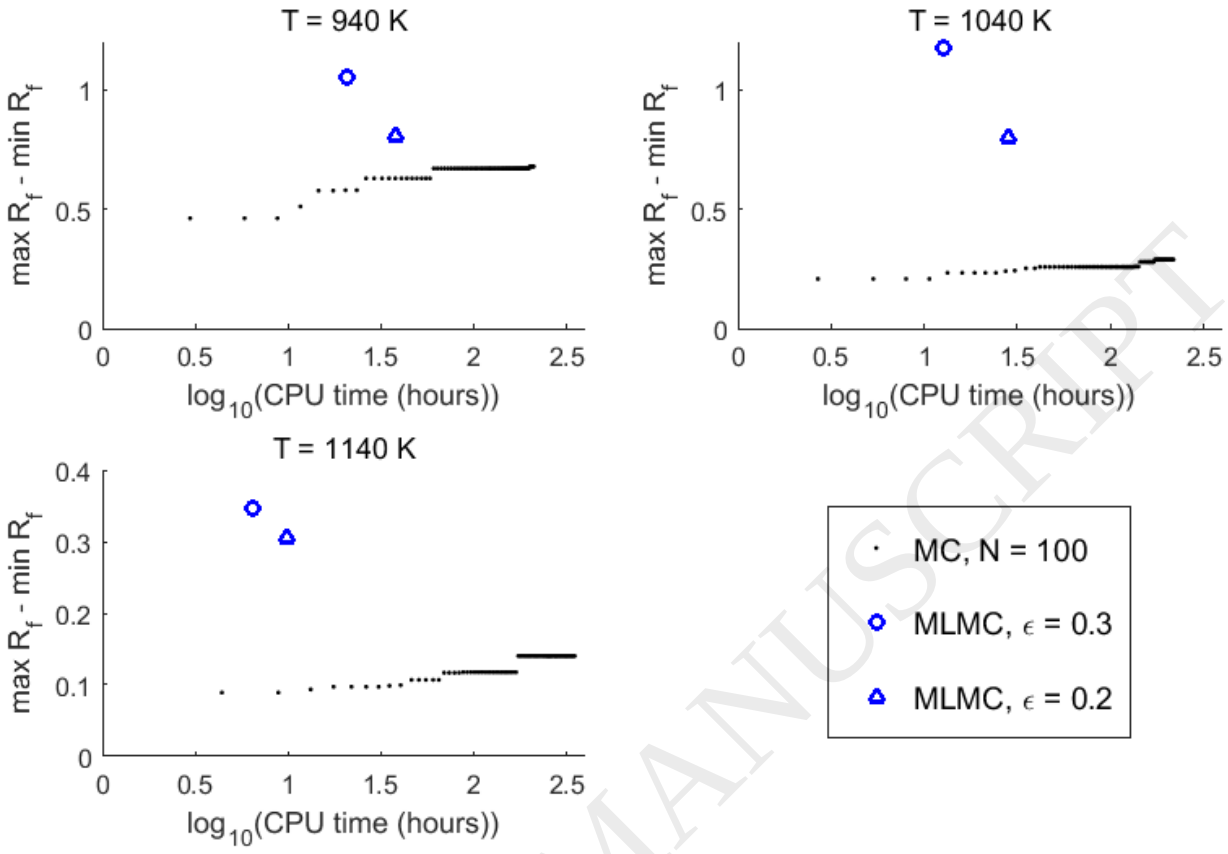


Figure 6: A comparison of the estimates of noise in R_f from MC and MLMC sampling.

Table 1: The parameters of the multiscale model

Symbol	Meaning
a	hydrodynamic strain rate
C_{tot}	concentration of sites on the film surface
E	energy associated with a single bond
E_d	energy required for desorption
E_m	energy required for migration
f	dimensionless stream function
i	number of neighbours of an atom (between one and five)
k_{d0}	event frequency constant
μ_b	viscosity of bulk gas
m	precursor molecular weight
M_i	number of adsorbed surface atoms with a particular number of neighbours (i)
η	dimensionless distance away from thin film surface
N	number of adsorption sites along an edge of the square thin film (set by the user)
N_a	number of adsorbed atoms
N_d	number of desorbed atoms
P	chamber pressure
Pr	Prandtl number
ρ	density of gas boundary layer
ρ_b	density of bulk gas
R_a	rate of adsorption from the gas boundary layer
R_d	rate of desorption to the gas boundary layer
R_g	gas constant
S_0	sticking coefficient
Sc	Schmidt number of the depositing gas species
τ	dimensionless time, equals to $2at$
t	time
Δt_{kMC}	kMC timestep
ΔT	kMC-PDE coupling time interval
T	substrate temperature (set by the user)
T_{bulk}	bulk temperature of the gas
$T_{surface}$	temperature of the substrate and the thin film
W_a	rate of adsorption onto the thin film
W_d	rate of desorption from the thin film
W_m	rate of migration on the thin film
x	mole fraction of the depositing gas species, varies with η and τ
X	mole fraction of the gas species in the bulk (i.e. at $\eta \to \infty$)
ς	uniform random number on the $(0,1)$ interval

Table 2: Summary of MC and MLMC results.

T *(K)*	*Sampling method*	*CPU time (hours)*	$max\ R_f - min\ R_f$
	MC, $N = 100$	208.3	0.679
940	MLMC, $\epsilon = 0.3$	21.0	1.054
	MLMC, $\epsilon = 0.2$	38.2	0.809
	MC, $N = 100$	217.6	0.289
1040	MLMC, $\epsilon = 0.3$	12.7	1.175
	MLMC, $\epsilon = 0.2$	28.5	0.803
	MC, $N = 100$	349.8	0.140
1140	MLMC, $\epsilon = 0.3$	6.5	0.347
	MLMC, $\epsilon = 0.2$	9.8	0.307