Available online at www.sciencedirect.com
ScienceDirect
Computers and Chemical Engineering 30 (2006) 1670–1686
Computers & Chemical Engineering
www.elsevier.com/locate/compchemeng

Control and optimization of multiscale process systems

Panagiotis D. Christofides [a,*], Antonios Armaou [b]

[a] Department of Chemical and Biomolecular Engineering, University of California, Los Angeles, CA 90095, United States
[b] Department of Chemical Engineering, The Pennsylvania State University, University Park, PA 16802, United States

Received 2 February 2006; received in revised form 9 May 2006; accepted 16 May 2006
Available online 14 July 2006

Abstract

In this work, we present an overview of recently developed methods for control and optimization of complex process systems described by multiscale models. We primarily discuss methods developed in the context of our previous research work and use examples of thin film growth processes to motivate the development of these methods and illustrate their application.

Keywords: Multiscale process systems; Kinetic Monte-Carlo simulation; Stochastic partial differential equation (sPDE); Feedback control; Thin film growth

1. Multiscale process systems: thin film growth

Over the last 10 years, increasingly tight product quality specifications have motivated extensive research on the development of control and optimization methods for distributed and multiscale process systems using increasingly detailed process descriptions. On one hand, for distributed process systems for which continuum laws are applicable, nonlinear distributed parameter systems, such as nonlinear hyperbolic/parabolic partial differential equations (PDEs), Navier–Stokes equations and population balance equations are employed for the design of high-performance feedback controllers used to regulate spatial temperature and concentration profiles in advanced materials processing applications, achieve wave suppression and drag reduction in fluid dynamic systems and shape particle size distribution in particulate processes, respectively (see, for example, the special issues Christofides, 2002b; Christofides & Armaou, 2005) and the books (Christofides, 2001, 2002a) for representative results and references in these areas). On the other hand, for processes that involve coupling of macroscale phenomena with important phenomena at mesoscopic/microscopic length scales, multiscale systems coupling continuum-type distributed parameter systems with molecular dynamics (MD) or kinetic Monte-Carlo (MC/kMC) simulations are employed because of their ability to describe phenomena which are inaccessible to continuum laws and equations.

An industrially important process where multiscale modeling is needed to adequately describe the coupling of macroscopic and microscopic phenomena is thin film growth. Thin films of advanced materials are currently used in a very wide range of applications, *e.g.*, microelectronic devices, optics, micro-electro-mechanical systems (MEMS) and biomedical products. Various deposition methods have been developed and widely used to prepare thin films such as physical vapor deposition (PVD) and chemical vapor deposition (CVD). However, the dependence of the thin film properties, such as uniformity, composition and microstructure, on the deposition conditions is a severe constraint on reproducing the thin film's performance. Thus, real-time feedback control of thin film deposition, based on fundamental models, becomes increasingly important in order to meet the stringent requirements on the quality of thin films and reduce thin film variability. While deposition, uniformity and composition control can be accomplished on the basis of continuum-type distributed parameter models (see, for example, Christofides, 2001; Theodoropoulou, Adomaitis, & Zafiriou, 1999 for results on rapid thermal processing (RTP) and Armaou & Christofides, 1999; Ni et al., 2004, on plasma-enhanced chemical vapor deposition (PECVD)), precise control of thin film microstructure requires multiscale distributed models that predict how the film state (microscopic scale) is affected by changes in the controllable process parameters (macroscopic scale).

* Corresponding author. Tel.: +1 310 794 1015; fax: +1 310 206 4107.
E-mail address: pdc@seas.ucla.edu (P.D. Christofides).

0098-1354/$ – see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.compchemeng.2006.05.025

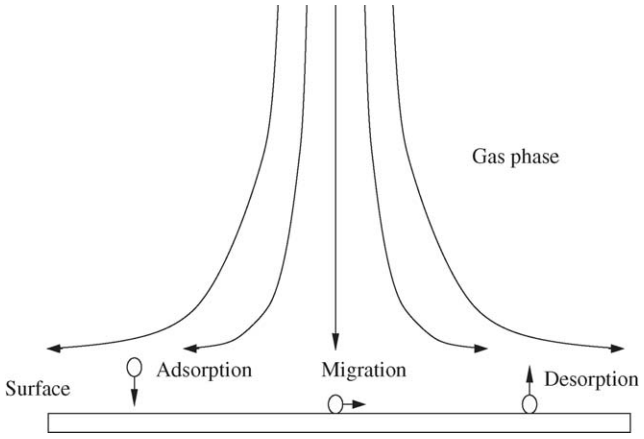

Fig. 1. Illustration of the thin film growth process.

In the remainder of this section, in order to discuss an example of multiscale modeling and provide the necessary background for presenting our methods for control and optimization of process systems using multiscale models, we consider the growth of a thin film from a fluid in a vertical, stagnation flow geometry. The process is shown in Fig. 1. In this geometry, inlet fluid flow forms a uniform boundary layer adjacent to the surface of the substrate and precursor atoms diffuse through such a boundary layer and deposit a thin film (Gadgil, 1993). Upon arrival at the surface, the precursor atoms are adsorbed onto the surface. Subsequently, adsorbed atoms may desorb to the gas phase or migrate on the surface.

From a modeling point of view, the major challenge is the integration of the wide range of length and time scales that the process encompasses (Vlachos, 1997). Specifically, in the gas phase, the processes of heat/mass transport can be adequately modeled under the hypothesis of continuum, thereby leading to PDE models for chamber temperature and species concentration. However, when the microstructure of the surface is studied, microscopic events such as atom adsorption, desorption and migration have to be considered, and the length scale of interest reduces dramatically to the order of that of several atoms. Under such a small length scale, the continuum hypothesis is no longer valid and deterministic PDEs cannot be used to describe the microscopic phenomena. Different approaches, such as Monte-Carlo simulation or molecular dynamics, should be employed to model the evolution of surface microstructure.

Although different modeling approaches are needed to model the macroscopic and microscopic phenomena of the process, there are strong interactions between the macroscale and microscale phenomena. For example, the concentration of the precursor in the inlet gas governs the rate of adsorption of atoms on the surface, which, in turn, influences the surface roughness. On the other hand, the density of the adatoms on the surface affects the rate of desorption of atoms from the surface to the gas phase, which, in turn, influences the gas phase concentration of the precursor. A multiscale model (Vlachos, 1997) is employed in this work to capture the evolution of both macroscopic and microscopic phenomena of the thin film growth process as well as their interactions. A set of PDEs derived from the mass, momentum and energy balances are used to describe the gas phase dynamics. Kinetic MC simulation is employed to capture the evolution of surface microstructure. Furthermore, the parameters of MC simulation such as the temperature and precursor concentration are provided by the solution of PDEs and the results from the kinetic MC simulation are used to determine the boundary conditions of the PDEs of the macroscopic model. In the remainder of this section, we describe the model for the gas phase and the surface microstructure for the thin film growth process of Fig. 1.

1.1. Gas phase model

Under the assumption of axisymmetric flow, the gas phase can be modeled through continuum type momentum, energy and mass balances as follows (Lam & Vlachos, 2001):

$$\frac{\partial}{\partial\tau}\left(\frac{\partial f}{\partial\eta}\right) = \frac{\partial^3 f}{\partial\eta^3} + f\frac{\partial^2 f}{\partial\eta^2} + \frac{1}{2}\left[\frac{\rho_b}{\rho} - \left(\frac{\partial f}{\partial\eta}\right)^2\right] \tag{1.1}$$

$$\frac{\partial T}{\partial\tau} = \frac{1}{Pr}\frac{\partial^2 T}{\partial\eta^2} + f\frac{\partial T}{\partial\eta} \tag{1.2}$$

$$\frac{\partial y_i}{\partial\tau} = \frac{1}{Sc_j}\frac{\partial^2 y_i}{\partial\eta^2} + f\frac{\partial y_i}{\partial\eta}. \tag{1.3}$$

The following boundary conditions are used for $\eta \to \infty$:

$$T = T_{bulk}, \quad \frac{\partial f}{\partial\eta} = 1, \quad y_j = y_{jb}, \quad j = 1, \ldots, N_g \tag{1.4}$$

and for $\eta \to 0$ (surface):

$$\begin{aligned} &T = T_{surface}, \quad f = 0, \quad \frac{\partial f}{\partial\eta} = 0 \\ &\frac{\partial y_j}{\partial\eta} = 0, \quad \text{for } j \neq \text{growing} \\ &\frac{\partial y_{growing}}{\partial\eta} = \frac{Sc_{growing}(r_a - r_d)}{\sqrt{2a\mu_b\rho_b}} \end{aligned} \tag{1.5}$$

where f is the dimensionless stream function, r_a and r_d are the rates of adsorption and desorption, respectively, η the dimensionless distance to the surface, ρ the density of the mixture, Pr the Prandtl number, y_j and Sc_j the mole fraction and Schmidt number of the species j, respectively, μ_b and ρ_b the viscosity and the density of the bulk, respectively, a the hydrodynamic strain rate and $\tau = 2at$ is the dimensionless time.

Although the macroscopic model describes the evolution of the precursor concentration and temperature (which influence the configuration of the growing surface), no direct information of the surface microstructure is available from the macroscopic model. Furthermore, the boundary conditions for the mass transfer equation of the growing species depend on the rate of adsorption and desorption. Therefore, a microscopic model is necessary to model the surface microstructure and to determine the boundary conditions of the mass transfer equation.

1.2. Surface microstructure model

The thin film growth of Fig. 1 includes three processes: the adsorption of atoms from the gas phase to the surface, the desorption of atoms from the surface to the gas phase and the migration of atoms on the surface. In this study, we consider multilayer growth and assume that all the surface sites are available for adsorption at all times, therefore, the adsorption rate is treated as site independent. For an ideal gas, the adsorption rate is given by the kinetic theory (Lam & Vlachos, 2001):

$$r_a = \frac{s_0 P}{\sqrt{2\pi m k T C_{tot}}} \tag{1.6}$$

where s_0 is the sticking coefficient, k the Boltzmann constant, P the partial pressure of the precursor, C_{tot} the concentration of sites on the surface, m the molecular weight of the precursor and T is the gas phase temperature above the surface. The rate of desorption of an atom depends on the atom's local micro-environment (*i.e.*, interactions with nearest neighbors) and the local activation energy. Under the consideration of only first nearest neighbor interactions, the desorption rate of an atom from the surface with n first nearest neighbors is:

$$r_d(n) = \nu_0 \exp\left(-\frac{nE}{kT}\right) \tag{1.7}$$

where E is the energy associated with a single bond on the surface and ν_0 is the frequency of events, which is determined by the following expression:

$$\nu_0 = k_{d0} \exp\left(-\frac{E_d}{kT}\right) \tag{1.8}$$

where k_{d0} is an event frequency constant and E_d is the energy associated with desorption. Finally, surface migration is modeled as desorption followed by re-adsorption (Gilmer & Bennema, 1972), and the migration rate is given by:

$$r_m(n) = \nu_0 A \exp\left(-\frac{nE}{kT}\right) \tag{1.9}$$

where A is associated with the energy difference that an atom on a flat surface has to overcome in jumping from one lattice site to an adjacent one and A is given as:

$$A = \exp\left(\frac{E_d - E_m}{kT}\right) \tag{1.10}$$

where E_m is the energy associated with migration.

The formation of the thin film by adsorption, migration and desorption is a stochastic process because: (a) the exact time and location of the occurrence of one specific surface micro-process (adsorption, migration or desorption) are unknown and (b) the probability with which each surface micro-process may occur is only available. Therefore, the surface evolution model should be established based on probability theory. Specifically, we treat the surface micro-processes as Poisson processes, which means that the following assumptions are made (Feller, 1975; Fichthorn & Weinberg, 1991; Gillespie, 1976; Melsa & Sage, 1973): (1) the probability that k events occur in the time interval $(t, t+T)$ is independent of t, (2) the probability that k events occur in the time interval $(t, t+T)$ is independent of the number of events occurring in any nonoverlapping time interval, and (3) the probability that an event occurs in an infinitesimal time interval $(t, t+\mathrm{d}t)$ is equal to $W\,\mathrm{d}t$ (where W is the mean count rate of the event), and the probability of more than one event occurring in an infinitesimal time interval is negligible.

Based on these three assumptions, the time evolution of probabilities that the surface is in one specific configuration can be derived. The configuration of a surface is characterized as the height of each surface atom at each surface site. If $P(\alpha, t)$ represents the probability that the system is in configuration α at time t, based on Assumptions 2 and 3 above, we have the following equation for $P(\alpha, t+\mathrm{d}t)$:

$$P(\alpha, t+\mathrm{d}t) = P(\alpha, t)P_{0\alpha} + \sum_{\beta} P(\beta, t)P_{1\beta} \tag{1.11}$$

where $P_{0\alpha}$ is the probability that no event occurs in the time interval $(t, t+\mathrm{d}t)$ given that the surface is in configuration α at t, $P(\beta, t)$ is the probability that the surface is in configuration β at t and $P_{1\beta}$ is the probability that one event occurs in the time interval $(t, t+\mathrm{d}t)$ given that the surface is in configuration β at t, and the occurrence of this event results to a transition from configuration β to configuration α. $P_{0\alpha}$ and $P_{1\beta}$ have the following expressions (a detailed proof can be found in Gillespie, 1992). Specifically,

$$P_{0\alpha} = 1 - \sum_{\beta} W_{\beta\alpha}\,\mathrm{d}t \tag{1.12}$$

where $W_{\beta\alpha}\,\mathrm{d}t$ is the probability that an event occurs in the time interval $(t, t+\mathrm{d}t)$ which results in a transition from configuration α to a configuration β, therefore, $\sum_{\beta} W_{\beta\alpha}\,\mathrm{d}t$ is the probability that any one event occurs in the time interval $(t, t+\mathrm{d}t)$ provided that the surface configuration is α at t. Moreover,

$$P_{1\beta} = W_{\alpha\beta}\,\mathrm{d}t \tag{1.13}$$

where $W_{\alpha\beta}\,\mathrm{d}t$ is the probability that an event happens in the time interval $(t, t+\mathrm{d}t)$ and the occurrence of this event results to a transition from configuration β to configuration α. By substituting Eqs. (1.12) and (1.13) into Eq. (1.11) and setting $\mathrm{d}t \to 0$, we obtain a differential equation describing the time evolution of the probability that the surface is in configuration α, Eq. (1.14):

$$\frac{\mathrm{d}P(\alpha, t)}{\mathrm{d}t} = \sum_{\beta} P(\beta, t)W_{\alpha\beta} - \sum_{\beta} P(\alpha, t)W_{\beta\alpha}. \tag{1.14}$$

Eq. (1.14) is the so-called "master equation" (ME) for a stochastic process. The ME has a simple, linear structure, however, it is difficult to write the explicit mathematical form of Eq. (1.14) for any realistic system because the number of the possible states is extremely large for most systems of a realistic size. For example, for a system with 10×10 sites and a maximum height of 1, the number of configurations is $2^{100} \approx 10^{30}$. This makes the direct solution of Eq. (1.14), for any system of meaningful size, using numerical methods for integration of ordinary differential equations (*e.g.*, Runge–Kutta) impossible.

Monte-Carlo techniques provide a way to obtain unbiased realizations of a stochastic process, which is consistent with the ME. The consistency of the Monte-Carlo simulation to the ME is based on the fact that in a Monte-Carlo simulation, a time sequence of Monte-Carlo events is constructed following a probability density function which is derived based on the same assumptions (Assumptions 1–3 above) as those used in the derivation of the master equation (Gillespie, 1976). A Monte-Carlo event is characterized by both the type of the event and the site in which the event is executed. We use $e(x; i, j)$ to represent a Monte-Carlo event of type x executed on the site (i, j) and $N \times N$ is the size of the lattice. The sequence of Monte-Carlo events can be constructed based on the probability density function, $F(\tau, e)$. In particular, $F(\tau, e)\mathrm{d}\tau$ is the probability at time t that event e will occur in the infinitesimal time interval $(t+\tau, t+\tau+\mathrm{d}\tau)$ and can be computed by:

$$F(\tau, e) = W_{\mathrm{e}} \exp(-W_{\mathrm{tot}}\tau). \tag{1.15}$$

Monte-Carlo simulation constructs the sequence of events following the probability density function shown in Eq. (1.15). There are many Monte-Carlo algorithms available to simulate a stochastic dynamic process. In the following calculations, the kinetic Monte-Carlo simulation algorithm developed by Vlachos (1997) is used. This algorithm is a modification of the so-called "direct" method developed by Gillespie (1976). Specifically, once the lattice is set and the probabilities of the three events are determined based on the corresponding rate expressions, a kinetic Monte-Carlo simulation is executed as follows: first, a random number is generated to select an event to be run based on the following probability:

$$P(e|\tau) = \frac{W_{\mathrm{e}}}{W_{\mathrm{tot}}} \tag{1.16}$$

then, a second random number is generated to select the site from the list of all available sites where the chosen event will be executed. This algorithm guarantees that every trial is successful and is efficient compared to traditional null event algorithms (Reese, Raimondeau, & Vlachos, 2001). Upon an executed event, a time increment $\mathrm{d}\tau$ is computed by Fichthorn and Weinberg (1991) and Lam and Vlachos (2001):

$$\mathrm{d}\tau = \frac{-\ln \xi}{r_{\mathrm{a}} \times N_{\mathrm{T}} + \nu_0(1+A)\sum_{m=1}^{5} N_m \exp\left(\frac{-mE}{kT}\right)} \tag{1.17}$$

where ξ is a random number in the (0,1) interval, N_{T} the total number of sites on the lattice and N_m is the number of atoms that have m neighbors on the surface.

In the remainder of this manuscript, we present an overview of recently developed methods for control and optimization of complex process systems on the basis of multiscale models. We will begin with control methods which directly utilize kinetic Monte-Carlo models, continue with results on the construction of closed-form stochastic PDEs from kinetic Monte-Carlo data and discuss their use in model-based predictive control and close with a method for optimization of process systems involving multiscale objectives.

2. Control using kinetic Monte-Carlo models

Methodologies for surface properties (*e.g.*, surface roughness) estimation and control using kinetic Monte-Carlo models have recently been developed in Lou and Christofides (2003a,b) and Ni and Christofides (2005b) and have been successfully applied (Lou & Christofides, 2004) to control surface roughness in a GaAs deposition process using experimentally determined model parameters. Furthermore, a method to construct reduced order approximations of the master equation was also reported in Gallivan and Murray (2004) and Gallivan et al. (2004). Surface roughness is a property of interest from a control point of view since it directly influences device properties. In this manuscript, the roughness, r, is represented by the number of broken bonds on the surface (Raimondeau & Vlachos, 2000):

$$r = \frac{\sum(|h_{i+1,j} - h_{i,j}| + |h_{i-1,j} - h_{i,j}| + |h_{i,j+1} - h_{i,j}| + |h_{i,j-1} - h_{i,j}|)}{2 \times N \times N} + 1 \tag{2.18}$$

where $N \times N$ is the dimension of the square lattice and $h_{i,j}$ is the number of atoms at site (i, j). For the thin film growth process of Fig. 1, the control problem is to achieve a desired surface roughness level by manipulating the substrate temperature. To address this problem, we need to develop an estimation scheme that will utilize roughness estimates obtained at discrete time instants (sampling times) to provide estimates of the surface roughness for all times. We present an estimation scheme which employs kinetic Monte-Carlo simulations of the surface together with roughness measurements obtained at discrete time instants to produce estimates of the surface roughness for all times. The basic idea is to construct a bank of 'parallel running' kinetic Monte-Carlo simulators of the surface based on small lattice size models to capture the dominant roughness evolution and utilize the available surface roughness measurements to improve upon the predictions of the kinetic Monte-Carlo simulators to obtain accurate surface roughness estimates.

Specifically, in the kinetic MC simulation, the size of the lattice influences the accuracy of the results and the computational demand. Roughly speaking, the computational complexity of the algorithm we adopt in this work is $O(N^4)$ and the magnitude of the fluctuation in the solution is $O(1/N^2)$ where N is the size of the lattice. The fourth-order dependence on computational complexity and the second-order dependence of fluctuations on the size of the lattice leave room for reducing the solution time with relatively small loss of accuracy. In our simulations, when the size of the lattice is reduced to 30×30, the solution time of the kinetic Monte-Carlo simulation is comparable to the real-time process evolution and the average values of the surface roughness approximate well the average values of these variables, which are obtained by running the kinetic Monte-Carlo simulation on a 120×120 lattice (this is a sufficiently large lattice to ensure simulation results which are independent of the lattice size).

However, the outputs from a kinetic Monte-Carlo simulation using a 30×30 lattice contain significant stochastic fluctuations, and thus, they cannot be directly used for feedback control (such

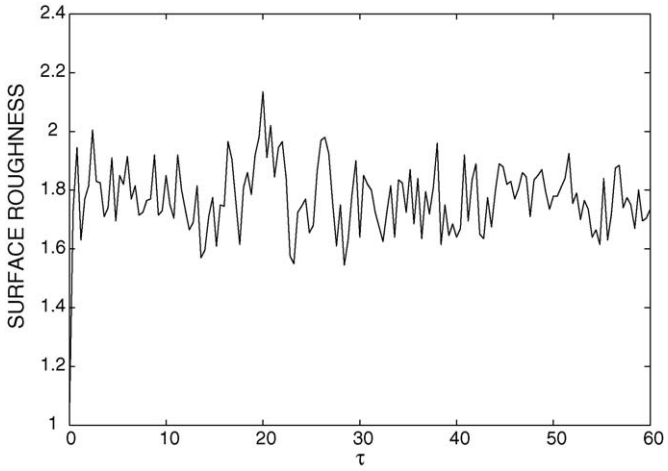

Fig. 2. Surface roughness from a kinetic Monte-Carlo simulation which uses a 20×20 lattice.

an approach would result to significant fluctuations of the control action which could perturb unmodeled (fast) process dynamics and should be avoided).

The fluctuations on the values of the outputs obtained from the kinetic Monte-Carlo simulation using the 30×30 lattice can be reduced by independently running several small lattice kinetic Monte-Carlo simulations with the same parameters and averaging the outputs of the different runs. Fig. 2 shows the surface roughness obtained from a Monte-Carlo simulation which uses a 20×20 lattice. Fig. 3 shows the surface roughness obtained from the computation of the average of six independent kinetic Monte-Carlo simulations which utilize a 20×20 lattice. These results show that when the outputs from multiple kinetic Monte-Carlo simulations that use small lattices are averaged, surface roughness fluctuations can be significantly reduced.

The predicted profiles of surface roughness, which are obtained from kinetic Monte-Carlo simulation based on multiple small lattice models, still contain stochastic fluctuations and are not robust (due to the open-loop nature of the calculation) with respect to disturbances and variations in process

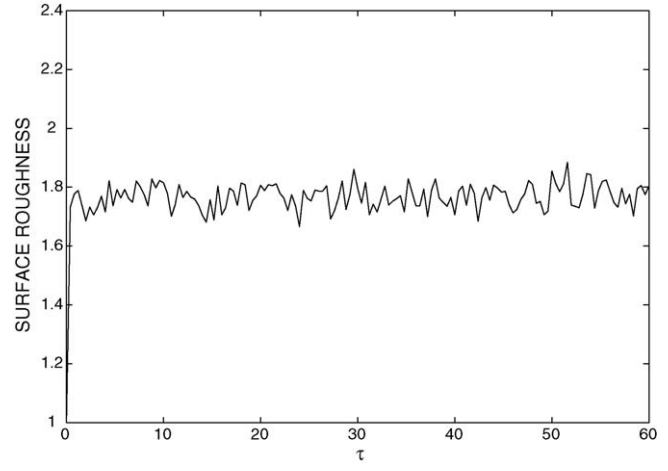

Fig. 3. Surface roughness from the computation of the average of six independent kinetic Monte-Carlo simulations which utilize a 20×20 lattice.

parameters. To alleviate these problems, we combine the small lattice kinetic Monte-Carlo simulators with an adaptive filter, to reject the stochastic fluctuations on the surface roughness and growth rate profiles, and a measurement error compensator to improve the estimates of these variables using on-line measurements. To simplify the notation of the mathematical formulas, we only present the general structure of the adaptive filter and of the measurement error compensator. Specifically, the adaptive filter is a second-order dynamical system with the following state-space representation:

$$\begin{aligned} \frac{\mathrm{d}\hat{y}_\mathrm{r}}{\mathrm{d}\tau} &= y_\mathrm{l} \\ \frac{\mathrm{d}y_\mathrm{l}}{\mathrm{d}\tau} &= \frac{K}{\tau_I}(y_\mathrm{r} - \hat{y}_\mathrm{r}) - \frac{1}{\tau_\mathrm{l}} y_\mathrm{l} \end{aligned} \tag{2.19}$$

where y_r is the output of the kinetic Monte-Carlo simulation based on multiple small lattice models, $\hat{y}_\mathrm{r}$ the filter output, K the filter gain and τ_I is the time constant. To accelerate the response of the filter and avoid large overshoot, $\tau_\mathrm{I} = 0.5/K$. To achieve both fast tracking of the dynamics of the outputs and efficient noise rejection, the gain of the filter is adaptively adjusted according to the following law:

$$K(\tau) = K_0 \frac{|f_{\tau-\Delta\tau}^{\tau} y_\mathrm{r}(t)\,\mathrm{d}t - f_{\tau-2\Delta\tau}^{\tau-\Delta\tau} y_\mathrm{r}(t)\,\mathrm{d}t|}{\Delta\tau^2} + K_\mathrm{s} \tag{2.20}$$

where K_0 is a constant, K_s the steady state gain for the adaptive filter and $\Delta\tau$ is the dimensionless time interval between two updates of K. Although a better tracking performance is expected when a small $\Delta\tau$ is used, a very small $\Delta\tau$ will introduce the effect of stochastic roughness fluctuations on the filter gain and estimates and should be avoided; the specific value of the appropriate $\Delta\tau$ is a function of the significance of the stochastic roughness fluctuations and its computation can be achieved through numerical simulation of the estimator for different values of $\Delta\tau$. The measurement error compensator uses the available on-line measurements to produce improved estimates of the surface roughness. The state-space representation of the measurement error compensator is:

$$\begin{aligned} &\frac{\mathrm{d}e}{\mathrm{d}\tau} = K_\mathrm{e}(y_\mathrm{h}(\tau_{m_i}) - \hat{y}(\tau_{m_i})); \quad \tau_{m_i} < \tau \leq \tau_{m_{i+1}} \\ &i = 1, 2, \ldots \end{aligned} \tag{2.21}$$

and the final roughness estimates are computed by:

$$\hat{y} = \hat{y}_\mathrm{r} + e. \tag{2.22}$$

In the above equations, K_e is the compensator gain, e the estimated model error, which is used to compensate the model output, $\hat{y}$ the roughness estimates, $\hat{y}_\mathrm{r}$ the filtered output from a kinetic Monte-Carlo simulator which uses a small lattice (or multiple small lattice models) and y_h is the output of a kinetic Monte-Carlo simulator which uses the large lattice (in an experimental set-up y_h could be obtained from the measurement sensor). Since the roughness measurements are only available at discrete points in time $\tau_m = [\tau_{m_1}, \tau_{m_2}, \ldots]$, the right-hand side of Eq. (2.21) is computed at the time a roughness measurement is available and is kept in this value in the time interval between two available roughness measurements.

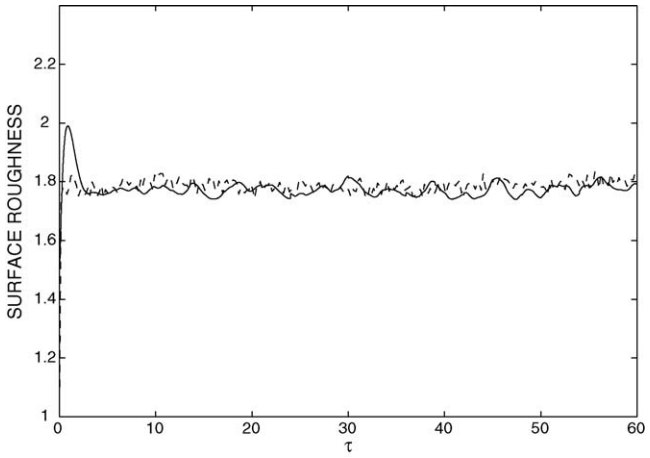

Fig. 4. Surface roughness profiles from the estimator (solid line) and from a kinetic Monte-Carlo simulation which uses a 120 × 120 lattice model (dashed line).

The combination of the adaptive filter and the measurement error compensator functions as an estimator, which is capable to accurately predict the evolution of surface roughness and growth rate during the thin film growth by using measurements of the precursor concentration above the substrate. In this work, we assume that measurements of precursor concentration above the substrate are available; when such measurements are not available, a state estimator can be constructed on the basis of the PDE model that describes the gas phase species concentrations and the temperature to obtain estimates of this quantity (for estimator design methods for PDE systems, see Christofides, 2001). Fig. 4 shows the surface roughness profile computed by the estimator, which uses a kinetic Monte-Carlo simulator based on six 20 × 20 lattice models (solid lines); it is compared with the surface roughness profile obtained from a kinetic Monte-Carlo simulator which uses a 120 × 120 lattice model. The sampling time $\tau_{m_{i+1}} - \tau_{m_i} = 3$. The results clearly show that the developed estimator can accurately predict the evolution of the surface roughness; this is the result of the use of kMC models in the estimation scheme coupled with measurements at distinct time instants to correct for model error. Note also that the developed estimator can be used for real-time feedback control since, the computational time needed to run kinetic Monte-Carlo simulation based on six 20 × 20 lattice models is comparable to the real-time process evolution. Finally, the reader may refer to Lou and Christofides (2003a) for simulation results that demonstrate that a proportional-integral (PI) controller that achieves very good performance on the basis of continuous roughness measurements exhibits very poor closed-loop performance when measurements at distinct time instants are used; in the same simulation set-up, the combination of the above estimation scheme with the PI controller allows achieving excellent closed-loop performance.

Referring to the selection of the lattice size, it is important to point out that while kinetic Monte-Carlo simulation based on multiple 20 × 20 lattice models can adequately capture the evolution of the surface roughness in the specific thin film growth problem under consideration, the dimension of the small lattice in general should be chosen so that the interactions between the surface atoms are adequately captured, and also that it is large enough to describe all the spatio-temporal phenomena occurring on the surface (*e.g.*, cluster formation). Furthermore, the small lattice should be chosen to provide accurate estimates of the desired properties to be controlled. For example, in the case of surface roughness, this quantity is defined as the average number of broken bonds for every surface atoms and the microscopic unit involved is an individual atom. When a small 20 × 20 lattice is used, the computation of surface roughness involves hundreds of surface atoms, which is adequate to obtain the expected value. However, when the property of interest is, for example, step density, a larger lattice is needed to obtain a convergent average value from the kinetic Monte-Carlo simulation.

We now turn our attention to the design and evaluation of a multivariable feedback control structure, based on kinetic Monte-Carlo models, used to control the surface roughness and growth in the thin film growth process of Fig. 1 by manipulating the substrate temperature and inlet precursor concentration. A diagram of the multivariable control system using the estimator/controller structure with interaction compensation is shown in Fig. 5. $G_1(s)$ is the transfer function between the substrate temperature and the growth rate and $G_2(s)$ is the transfer function between the inlet precursor mole fraction and the growth rate. Step tests were used to identify the expression and parameters of $G_1(s)$ and $G_2(s)$.

A closed-loop system simulation is performed to evaluate the effectiveness of the multivariable estimator/control structure with interaction compensation. Initially, the substrate temperature is $T = 800$ K and the inlet precursor mole fraction is 2.0×10^{-5}; these conditions correspond to a growth rate of about 180 ML/s and a surface roughness of about 1.8. The proposed multivariable control system (Fig. 5) is applied to the process to regulate the growth rate and surface roughness to the desired set-point values. The controller successfully drives both the surface roughness and the growth rate to the desired set-point values; the reader may refer to Lou and Christofides (2003b) for a detailed simulation study. Fig. 6 shows the surface roughness under multivariable control with interaction compensation.

3. Control using stochastic PDEs

While it is possible in certain examples to use kinetic Monte-Carlo models for real-time estimation and control, there are many applications where closed-form models are needed, owing to their computational efficiency, to carry out system-level analysis as well as design and implementation of real-time model-based feedback control systems. Motivated by this, several methods were reported in Armaou, Siettos, & Kevrekidis (2004), Armaou, Kevrekidis, & Theodoropoulos (2005), Mastny et al. (2005) and Siettos et al. (2003) to identify linear deterministic models from outputs of kinetic Monte-Carlo simulators which were subsequently used to design controllers using linear control theory to control macroscopic variables that are low statistical moments of the microscopic distributions (*e.g.*, surface coverage, which is the zeroth moment of adspecies distribution on a lattice). This approach was recently extended to the construc-

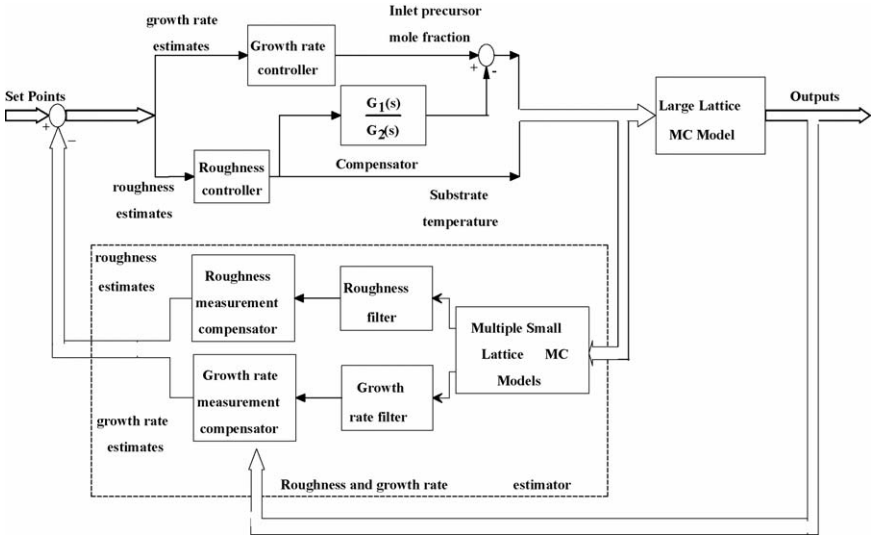

Fig. 5. Diagram of multivariable feedback control system with interaction compensation.

tion of nonlinear models (Armaou, 2005) and the subsequent design of nonlinear controllers (Armaou, 2005; Siettos et al., 2005). In this direction, other results also include the construction of linear/nonlinear deterministic models from input/output data using system identification techniques (Drew et al., 2004; Rusli et al., 2005; Wolfrum et al., 2005). However, to control higher statistical moments of the microscopic distributions, such as the surface roughness (the second moment of height distribution on a lattice), or even the microscopic configuration (such as the surface morphology), linear or nonlinear deterministic models may not be sufficient, because the effect of the stochastic nature of the microscopic processes becomes very significant and must be addressed both in the model construction and controller design. In such a case, stochastic differential equation models should be used. It turns out that there is a significant body of literature in the area of statistical physics (*e.g.*, Marsili, Maritan, Toigo, & Banavar, 1996; Park, Kim, & Park, 2002; Vvedensky, 2003) focusing on the construction of stochastic PDE models for thin film growth processes by appropriately averaging microscopic process rules in the context of discrete lattice models. Specifically, stochastic PDE models have been developed to describe the evolution of the height profile for surfaces in certain physical and chemical processes such as epitaxial growth (Vvedensky, Zangwill, Luse, & Wilby, 1993) and ion sputtering (Lauritsen, Cuerno, & Makse, 2003). Taking advantage of these results, we (Lou & Christofides, 2005b) presented a method for feedback control of surface roughness in a thin film growth process whose surface height fluctuation can be described by the Edwards–Wilkinson equation (Edwards & Wilkinson, 1982), a second-order stochastic parabolic PDE (see also Lou & Christofides, 2005a, for results on linear covariance control of surface roughness in a sputtering process using the stochastic Kuramoto–Sivashinsky equation). Specifically, a feedback controller was designed based on the stochastic PDE model and successfully applied to the kMC model of the deposition process regulating the surface roughness to desired values. However, the construction of stochastic PDE models for thin film growth processes directly based on microscopic process rules is a very difficult task. This issue has prohibited the development of stochastic PDE models, and subsequently the design of model-based feedback control systems, for realistic deposition processes which are, in general, highly complex.

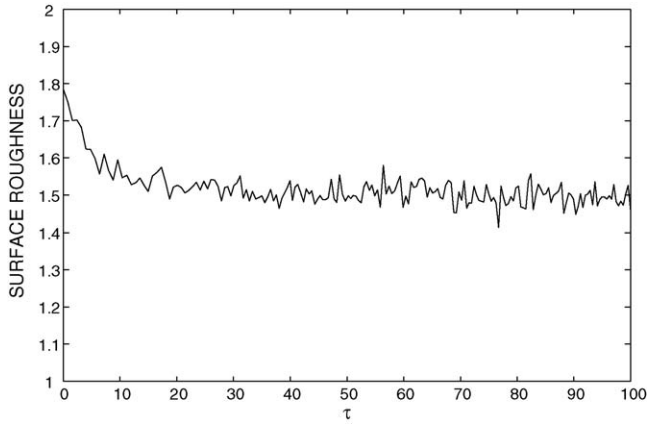

Fig. 6. Closed-loop surface roughness under multivariable feedback control—surface roughness set-point value is 1.5.

3.1. Construction of stochastic PDEs

Motivated by this practical problem, we recently presented (Ni & Christofides, 2005a) a systematic method for the construction of linear stochastic PDE models for feedback control of surface roughness in thin film deposition. To present the essential components of this method, we focus on a thin film growth process taking place on a one-dimensional lattice (extensions to two-dimensions are notationally involved and can be found in Ni & Christofides (2005c)). Without any *a priori* knowledge of the deposition process, we assume that there exists a one-dimensional linear stochastic PDE of the following general form that can adequately describe the evolution of the surface of the

thin film during the deposition:

$$\frac{\partial h}{\partial t} = c + c_0 h + c_1 \frac{\partial h}{\partial x} + c_2 \frac{\partial^2 h}{\partial x^2} + \cdots + c_w \frac{\partial^w h}{\partial x^w} + \xi(x,t) \tag{3.23}$$

where $x \in [0,\pi]$ is the spatial coordinate, t is the time, $h(x, t)$ is the height of the surface at position x and time t, and $\xi(x, t)$ is a Gaussian noise with zero mean and covariance:

$$\langle \xi(x,t)\xi(x',t') \rangle = \varsigma^2 \delta(x-x')\delta(t-t') \tag{3.24}$$

where $\delta(\cdot)$ is the Dirac function. Furthermore, the coefficients c and c_j in Eq. (3.23) and the parameter ς^2 in Eq. (3.24) depend on the process parameters (gas flow rates, substrate temperature, *etc.*) $p_i(t)$:

$$\begin{aligned} c &= C[p_1(t), p_2(t), \ldots, p_d(t)] \\ c_j &= C_j[p_1(t), p_2(t), \ldots, p_d(t)], \quad j = 0, \ldots, w \\ \varsigma^2 &= C_\xi[p_1(t), p_2(t), \ldots, p_d(t)] \end{aligned} \tag{3.25}$$

where $C(\cdot)$, $C_j(\cdot)$ and $C_\xi(\cdot)$ are nonlinear functions to be determined.

The stochastic PDE of Eq. (3.23) is subjected to the following periodic boundary conditions:

$$\frac{\partial^j h}{\partial x^j}(0,t) = \frac{\partial^j h}{\partial x^j}(\pi,t), \quad j = 0, \ldots, w-1 \tag{3.26}$$

and the initial condition:

$$h(x,0) = h_0(x) \tag{3.27}$$

To study the dynamics of Eq. (3.23), we initially consider the eigenvalue problem of the linear operator of Eq. (3.23), which takes the form:

$$\begin{aligned} A\phi_n(x) &= c_0\phi_n(x) + c_1 \frac{\mathrm{d}\phi_n(x)}{\mathrm{d}x} + c_2 \frac{d^2\phi_n(x)}{\mathrm{d}x^2} \\ &\quad + \cdots + c_w \frac{\mathrm{d}^w \phi_n(x)}{\mathrm{d}x^w} = \lambda_n \phi_n(x) \\ \frac{\mathrm{d}^j \phi_n}{\mathrm{d}x^j}(0) &= \frac{\mathrm{d}^j \phi_n}{\mathrm{d}x^j}(\pi), \quad j = 0, \ldots, w-1, \\ n &= 1, \ldots, \infty \end{aligned} \tag{3.28}$$

where λ_n denotes an eigenvalue and ϕ_n denotes an eigenfunction. A direct computation of the solution of the above eigenvalue problem yields:

$$\begin{aligned} \lambda_n &= c_0 + I2nc_1 + (I2n)^2 c_2 + \cdots + (I2n)^w c_w \\ \phi_n(x) &= \sqrt{\frac{1}{\pi}}\, \mathrm{e}^{I2nx}, \quad n = 0, \pm 1, \ldots, \pm\infty \end{aligned} \tag{3.29}$$

where λ_n denotes the nth eigenvalue, $\phi_n(x)$ denotes the nth eigenfunction and $I = \sqrt{-1}$.

To present the method that we use for parameter identification of the stochastic PDE of Eq. (3.23), we first derive an infinite-order stochastic ODE representation of Eq. (3.23) using modal decomposition and parameterize the infinite-order stochastic ODE system using kMC simulation. We first expand the solution of Eq. (3.23) in an infinite series in terms of the eigenfunctions of the operator of Eq. (3.28) as follows (*i.e.*, the Fourier expansion in the complex form):

$$h(x,t) = \sum_{n=-\infty}^{\infty} z_n(t)\phi_n(x) \tag{3.30}$$

where $z_n(t)$ are time-varying coefficients. Substituting the above expansion for the solution, $h(x, t)$, into Eq. (3.23) and taking the inner product, the following system of infinite stochastic ODEs is obtained:

$$\frac{\mathrm{d}z_n}{\mathrm{d}t} = \lambda_n z_n + c_{zn} + \xi_n(t), \quad n = 0, \pm 1, \ldots, \pm\infty \tag{3.31}$$

and the initial conditions:

$$z_n(0) = z_{n0}, \quad n = 0, \pm 1, \ldots, \pm\infty \tag{3.32}$$

where $c_{zn} = c\int_0^\pi \phi_n(x)\,\mathrm{d}x$ (apparently $c_{z0} = c\sqrt{\pi}$ and $c_{zn} = 0\ \forall\ n \neq 0$), $\xi_n(t) = \int_0^\pi \xi(x,t)\phi_n(x)\,\mathrm{d}x$ and $z_{n0} = \int_0^\pi h_0(x)\phi_n(x)\,\mathrm{d}x$.

The covariances of $\xi_n(t)$ can be computed as $\langle \xi_n(t) \rangle = 0$ and $\langle \xi_n(t)\xi_n^*(t') \rangle = \varsigma^2\delta(t-t')$ (ξ_n^* is the complex conjugate of ξ_n, the superscript star is used to denote complex conjugate in the remainder of this manuscript). We note that $\xi_n(t)$ is a complex Gaussian random variable and the probability distribution function of the Gaussian distribution, $P(\xi_n, t)$, on the complex plane with zero mean and covariance $\varsigma^2\delta(t-t')$ is defined as follows:

$$P(\xi_n,t) = \frac{1}{\sqrt{2\pi}\varsigma\delta(t-t')} e^{\xi_n\xi_n^*/2\varsigma^2\delta(t-t')} \tag{3.33}$$

To parameterize this system of infinite stochastic ODEs, we first derive the analytic expressions for the statistical moments of the stochastic ODE states, including the expected value and covariance. By comparing the analytical expression to the statistical moments obtained by multiple kMC simulations, the parameters of the stochastic ODE system (*i.e.*, λ_n and ς) can be determined.

The analytic solution of Eq. (3.31) is obtained as follows to derive the expressions for the statistical moments of the stochastic ODE states:

$$z_n(t) = \mathrm{e}^{\lambda_n t} z_{n0} + \frac{(\mathrm{e}^{\lambda_n t} - 1)c_{zn}}{\lambda_n} + \theta_n(t) \tag{3.34}$$

where $\theta_n(t)$ is a complex random variable of normal distribution with zero mean and covariance $\langle \theta_n(t)\theta_n^*(t) \rangle = \varsigma^2(\mathrm{e}^{(\lambda_n+\lambda_n^*)t} - 1/\lambda_n + \lambda_n^*)$. Therefore, the expected value (the first stochastic moment) and the covariance (the second stochastic moment) of state z_n can be expressed as follows:

$$\begin{aligned} \langle z_n(t) \rangle &= \mathrm{e}^{\lambda_n t} z_{n0} + \frac{(\mathrm{e}^{\lambda_n t} - 1)c_{zn}}{\lambda_n} \\ \langle z_n(t) z_n^*(t) \rangle &= \varsigma^2 \frac{\mathrm{e}^{(\lambda_n+\lambda_n^*)t} - 1}{\lambda_n + \lambda_n^*} + \langle z_n(t) \rangle \langle z_n(t) \rangle^* \\ n &= 0, \pm 1, \ldots, \pm\infty. \end{aligned} \tag{3.35}$$

Eq. (3.35) holds for any initial condition z_{n0}. Since we are able to choose any initial thin film surface for simulation, we choose $z_{n0} = 0$ (*i.e.*, the initial surface is flat, $h(x, 0) = 0$) to simplify our

calculations. In this case, Eq. (3.35) can be further simplified as follows (note that $c_{zn} = 0$, $\forall n \neq 0$):

$$\langle z_n(t) \rangle = 0$$
$$\left\langle z_n(t) z_n^*(t) \right\rangle = \varsigma^2 \frac{e^{(\lambda_n + \lambda_n^*)t} - 1}{\lambda_n + \lambda_n^*} = \varsigma^2 \frac{e^{2Re(\lambda_n)t} - 1}{2Re(\lambda_n)} \tag{3.36}$$
$$n = \pm 1, \ldots, \pm\infty$$

where $\mathrm{Re}(\lambda_n)$ denote the real part of λ_n, and for $z_0(t)$, it follows from Eq. (3.35) with $\lambda_0 = 0$ that

$$\langle z_0(t) \rangle = \lim_{\lambda_0 \to 0} \frac{(e^{\lambda_0 t} - 1) c_{z0}}{\lambda_0} = t c_{z0} = t\sqrt{\pi} c$$
$$\left\langle z_0^2(t) \right\rangle = \varsigma^2 t + t^2 \pi c^2. \tag{3.37}$$

It can be seen in Eq. (3.36) that the statistical moments of each stochastic ODE state depend only on the real part of the corresponding eigenvalue, and therefore, to determine the imaginary part of the eigenvalue we construct an extra equation related to the expected value of $\mathrm{Re}[\lambda_n(t)]^2$ (not shown here due to space limitations). We note that λ_n would be a complex number if the linear operator A is not self-adjoint i.e., for example, when odd-partial-derivatives are present in the stochastic PDE see (Eq. (3.29)).

Based on the above results, we proposed a systematic procedure to construct linear stochastic PDEs for the deposition process based on a kinetic Monte-Carlo code used to simulate the deposition process and generate surface snapshots. The proposed procedure includes the following steps: first, we design a set of simulation experiments that cover the complete range of process operation; second, we run multiple simulations for each simulation experiment to obtain the trajectories of the first and second statistical moments of the states (*i.e.*, Fourier coefficients) computed from the surface snapshots; third, we compute the eigenvalues of the linear operator and covariance of the Gaussian noise based on the trajectories of the statistical moments of the states for each simulation experiment, and determine the model parameters of the stochastic PDE (*i.e.*, the pre-derivative coefficients and the order of the stochastic PDE); finally, we investigate the dependence of the model parameters of the stochastic PDE on the process parameters and determine the least-square-optimal form of the stochastic PDE model with model parameters expressed as functions of the process parameters.

Because there are only two process parameters considered in the deposition process studied in this work, the growth rate W and the substrate temperature T, the simulation experiment design is straightforward. Specifically, different W values and T values are evenly selected from the range of process operation of interest and simulation experiments are executed with every selected W value for each selected T value. Therefore, we start our demonstration of the model construction methodology with the identification of the eigenvalues and covariance. Also, we note that the trajectories of the statistical moments for each simulation experiment are computed based on 100 simulation runs taking place with the same process parameters.

Since for a flat initial surface, the covariance of each state $\left\langle z_n(t) z_n^*(t) \right\rangle$ should be able to be predicted by Eq. (3.36), therefore, we can fit ς^2 and $\mathrm{Re}(\lambda_n)$ in Eq. (3.36) for the profile of $\left\langle z_n(t) z_n^*(t) \right\rangle$. In order to obtain the profile of $\left\langle z_n(t) z_n^*(t) \right\rangle$ we need to generate snapshots of the thin film surface during each deposition simulation and compute the values of $z_n(t)$. Since the lattice consists of discrete sites, we let $h(kL, t)$ be the height profile of the surface at time t with lattice constant L (k denotes the coordinate of a specific surface site), and compute $z_n(t)$ as follows:

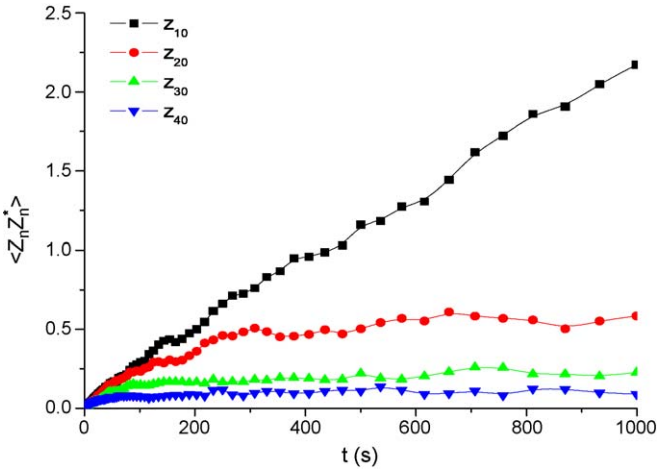

Fig. 7. Covariance profiles of z_{10}, z_{20}, z_{30} and z_{40}.

$$z_n(t) = \sum_{k=0}^{k_{\max}} h(kL, t) \int_{kL}^{(k+1)L} \phi_n^*(x)\, dx \tag{3.38}$$

where $k_{\max} L = \pi$ (*i.e.*, the lattice is mapped to the domain $[0,\pi]$). Substituting Eq. (3.29) into Eq. (3.38), we can derive the following expression for $z_n(t)$:

$$z_n(t) = \sum_{k=0}^{k_{\max}} \frac{h(kL, t)\, e^{-2kLnI}}{2\sqrt{\pi} nI} (1 - e^{-2LnI}) \tag{3.39}$$
$$n = \pm 1, \ldots, \pm\infty$$

and for $z_0(t)$, we have,

$$z_0(t) = \sum_{k=0}^{k_{\max}} h(kL, t) \frac{L}{\sqrt{\pi}} = t\sqrt{\pi} W. \tag{3.40}$$

Fig. 7 shows the typical covariance profiles of different states in a growth process. It can be seen that despite the very different time scales of the states, our method can still generate very smooth profiles for both the fast states (such as z_{40}, whose time scale is less than 50 s) and the slow states (such as z_{10}, whose time scale is larger than 1000 s).

Fig. 8 shows the eigenvalues identified from thin film depositions occurring under the same operating conditions but simulated with different lattice size (we note that the identified eigenvalues are considered real since the imaginary part of the eigenvalues identified turned out to be very small). It can be seen that the identified spectra are very close to each other when n is rescaled with the corresponding lattice size. This is expected, since, $\phi_n(x)$ is a basis of the domain of operator A, and is a complex function of the frequency n. Accordingly, $n/k_{\max}$ is the

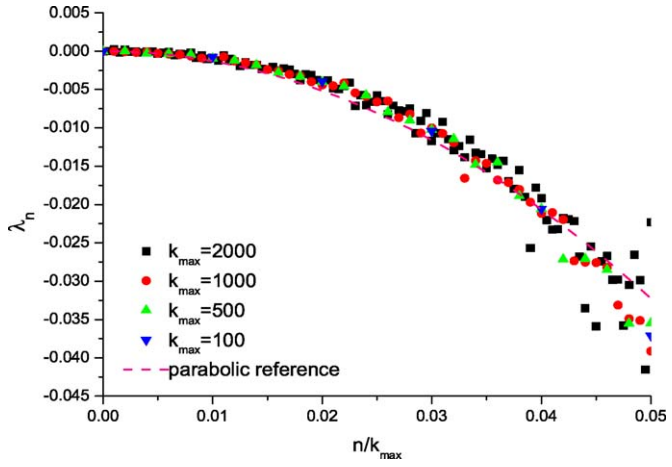

Fig. 8. Eigenvalue spectra of the infinite stochastic ODE systems identified from the kMC simulation of the deposition process with different lattice size: $k_{max} = 100$, 500, 1000 and 2000.

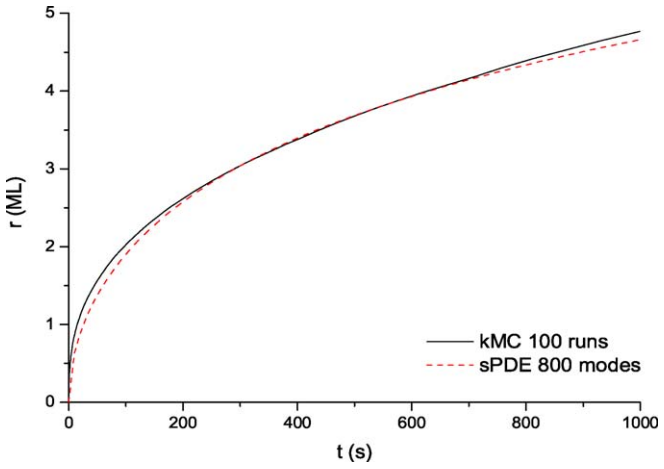

Fig. 9. Expected surface roughness profiles generated by kMC simulation and stochastic PDE model for a 1000 s deposition with substrate temperature $T = 550$ K, thin film growth rate $W = 0.1$ monolayer s^{-1} and lattice size $k_{max} = 2000$.

length scale of the surface fluctuation described by $\phi_n(x)$ when a lattice of size k_{max} is mapped to the domain of $[0,\pi]$ (we note that, for the same reason, the covariance values should be scaled with the inverse of the lattice size, $1/k_{max}$, in order to carry out a meaningful comparison).

It can also be seen in Fig. 8 that the eigenspectra are very close to the parabolic reference curve, which implies that a second-order stochastic PDE system of the following form would be able to describe the evolution of the surface height of this deposition process:

$$\frac{\partial h}{\partial t} = c + c_2 \frac{\partial^2 h}{\partial x^2} + \xi(x,t) \tag{3.41}$$

in which c, c_2 and the covariance of the Gaussian noise ξ, ς, all depend on the microscopic processes and operating conditions.

We proceed now with the derivation of the parameters of the stochastic PDE of Eq. (3.41). From Eqs. (3.37) and (3.40), we can see that $c = W$ for all cases. However, c_2 and ς^2 identified for different deposition settings can be very different, therefore, we need to investigate their dependence on the deposition parameters to obtain their empirical explicit expressions. c_2 and ς^2 are evaluated for assorted deposition conditions and a lattice size of 1000 (*i.e.*, $k_{max} = 1000$) is used for all the simulation runs in our study. To derive explicit expressions for c_2 and ς^2 as functions of T and W, we evaluate these values for different T and W and the results are shown below:

$$c_2(T,W) = \frac{e^{-32.002+0.0511T-0.1620W}}{k_{max}^2}$$
$$\varsigma^2(T,W) = 5.137 \times 10^{-8} T + 3.2003 \times 10^{-3} W. \tag{3.42}$$

Therefore, the linear stochastic PDE model identified for the deposition process is as follows:

$$\frac{\partial h}{\partial t} = W + \left(\frac{e^{-32.002+0.0511T-0.1620W}}{k_{max}^2} \right) \frac{\partial^2 h}{\partial x^2} + \xi(x,t);$$
$$\frac{\partial h}{\partial x}(0,t) = \frac{\partial h}{\partial x}(\pi,t), \quad h(0,t) = h(\pi,t) \tag{3.43}$$

where $\langle \xi(x,t)\xi(x',t') \rangle = (5.137 \times 10^{-8} T + 3.2003 \times 10^{-3} W)\delta(x-x')\delta(t-t')$.

We now proceed with the validation of the stochastic PDE model of the thin film deposition process (Eq. (3.43)). Validation experiments are conducted for a number of deposition conditions which have not been used for the model construction.

Fig. 9 shows the expected roughness profile of a deposition with substrate temperature $T = 550$ K and thin film growth rate $W = 0.1$ monolayer s^{-1}; Fig. 10 shows the roughness profile of a deposition with substrate temperature $T = 700$ K and thin film growth rate $W = 2.5$ monolayer s^{-1}; we can see that the linear stochastic PDE model constructed for the deposition process is also very consistent with the kinetic Monte-Carlo simulation in terms of surface roughness, at both low and high substrate temperatures, for different growth rates. The reader may refer to Ni and Christofides (2006) for additional model validation results. Of course, the ultimate test of the validity of the model

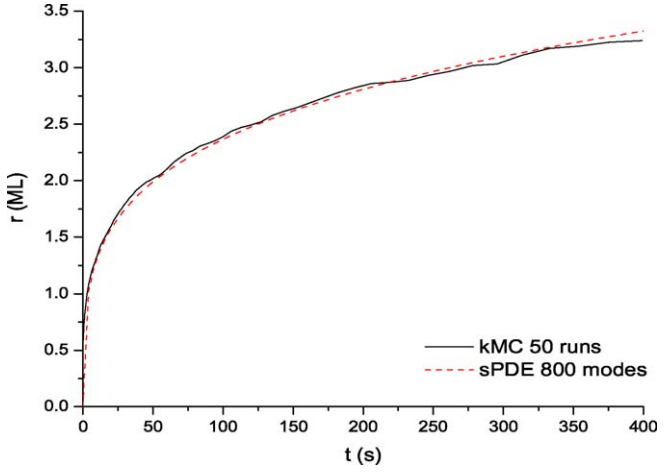

Fig. 10. Expected surface roughness profiles generated by kMC simulation and stochastic PDE model for a 400 s deposition with substrate temperature $T = 700$ K, thin film growth rate $W = 2.5$ monolayer s^{-1} and lattice size $k_{max} = 2000$.

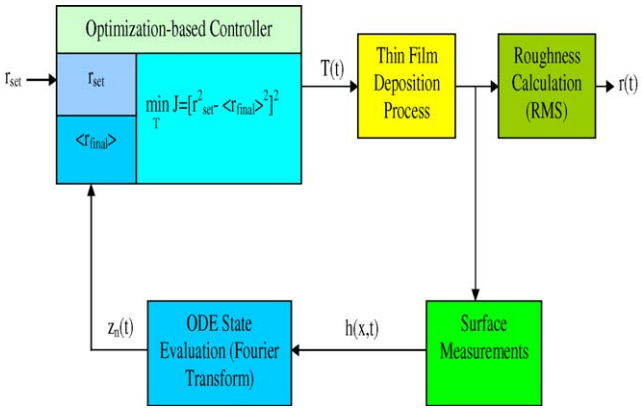

Fig. 11. Block diagram of the closed-loop system.

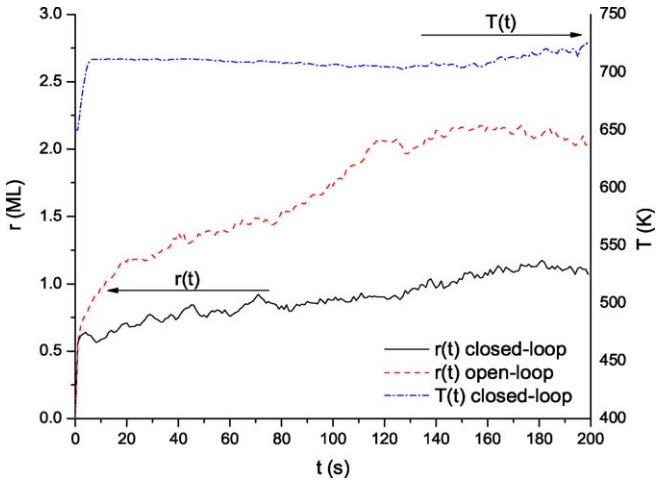

Fig. 12. Surface roughness and substrate temperature profiles of a 200 s closed-loop deposition process with thin film growth rate $W = 0.5$ monolayer s^{-1} and final roughness setpoint $r_{set} = 1.0$ monolayer.

is the closed-loop performance achieved by the controller that is designed on the basis of this model and this is discussed in the next subsection.

3.2. Predictive control using stochastic PDEs

We now proceed with the design of the feedback controller. Since the thin film deposition is a batch process, the control objective is to control the final surface roughness of the thin film to a desired level at the end of each deposition run by explicitly accounting for the presence of constraints on the manipulated input (an important problem that cannot be addressed by classical (PID) control schemes). Therefore, we use an optimization-based control problem formulation. Fig. 11 shows the block diagram of the closed-loop system. When a real-time surface profile measurement is obtained, the states of the infinite stochastic ODE system, z_n, are computed. Then, a substrate temperature T is computed based on states z_n and the stochastic PDE model and applied to the deposition process. The substrate is held at this temperature for the rest of the deposition until a different value is assigned by the controller. The value of T is determined at each time t by solving, in real-time, an optimization problem minimizing the difference between the estimated final surface roughness and the desired level.

A standard procedure based on the active set method is used to solve the optimization problem. A kMC code with a lattice size $k_{max} = 1000$ is used to simulate the thin film deposition process, and the substrate temperature is restricted within 300–900 K. The measurement interval, as well as the control interval, is set to be 1 s. We limit the maximum number of states to be used (in our case, to $m = 500$) to guarantee the maximum possible computation time for each control action is within certain requirement, however, for most of the time the number of states needed by the controller is much smaller.

Fig. 12 shows the surface roughness and substrate temperature profiles of a closed-loop deposition process with thin film growth rate $W = 0.5$ monolayer s^{-1}. The control objective is to drive the final surface roughness of the thin film to 1.0 monolayer at the end of a 200 s deposition. It can be seen that the final surface roughness is controlled at the desired level while an open-loop deposition with the same initial deposition condition would lead to a 100% higher final surface roughness as shown in Fig. 12. It is important to note that while the closed-loop profiles of Fig. 12 have been obtained under the assumption of availability of surface height profile measurements (state feedback control), the stochastic PDE model of Eq. (3.43) can be used as the basis for the design of a state estimator that employs on-line surface height measurements at specific spatial points to estimate the entire surface height profile and the corresponding roughness.

4. Optimization using multiscale models

4.1. Problem formulation and solution method

In this section, we discuss a method (Varshney & Armaou, 2005a,b) for optimization of process systems involving multiscale optimization objectives. Consider a process whose domain Ω comprises of a subdomain Ω_1 in which a closed-form macroscopic process description is available and a subdomain Ω_2 in which a microscopic process description is available. Mathematically, the multiscale process can be represented as:

$$0 = A(x) + f(x, d), \quad \text{on } \Omega_1$$
$$d = \sum_{i}^{n} d_i(z)(H(t - \bar{t}_i) - H(t - \bar{t}_{i+1})) \tag{4.44}$$

$$x_m(t_i) = \Pi(x_m(t_{i-1}), \delta t, x|_\gamma), \quad \text{on } \Omega_2$$
$$\delta t = t_i - t_{i-1} \tag{4.45}$$

$$g\left(x, \frac{\mathrm{d}x}{\mathrm{d}\eta}\right) = 0, \quad \text{on } \Gamma \backslash \gamma \tag{4.46}$$

$$h\left(\bar{x}_s, x|_\gamma, \frac{\mathrm{d}x}{\mathrm{d}\eta}\right) = 0, \quad \text{on } \gamma. \tag{4.47}$$

Eqs. (4.44) and (4.45) represent the macroscopic and microscopic descriptions of the process over the respective domains Ω_1 and Ω_2. It is assumed that Ω_1 and Ω_2 do not overlap and

share a common interface γ, and $\Omega = \Omega_1 \cup \Omega_2$ spans the whole process domain. $x(z) \in \mathbb{R}^N$ denotes the vector of macroscopic state variables, $x_m(t_i)$ is the vector of microscopic state variables at time-instant t_i, $z = [z_1, z_2, z_3] \in \Omega_1 \subset \mathbb{R}^3$ is the vector of spatial coordinates and Γ is the boundary of the macroscopic domain Ω_1. $A(x)$ is a second-order dissipative, possibly nonlinear, spatial differential operator, $f(x, d)$ is a nonlinear vector function which is assumed to be sufficiently smooth with respect to its arguments, $d \in \mathbb{R}^p$ is the vector of design variables and $\bar{t}_i$ is the time-instant when the design variables are varied. $g(x(\mathrm{d}x/\mathrm{d}\eta))$, defined on the boundary $\Gamma\backslash\gamma$ is a nonlinear vector function which is assumed to be sufficiently smooth, and η is the spatial direction perpendicular to the boundary Γ. It is assumed that the time horizon over which all the dynamics of the eigenmodes of Eq. (4.44) relax, t_i^s, is negligible in comparison to $\delta\bar{t}_i = \bar{t}_{i+1} - \bar{t}$, implying that the process in Ω_1 is operating under quasi-steady-state conditions. Function, Π, can be thought of as a black-box timestepper, which interacts with the macroscopic process model via an input/output structure and may be unknown in closed-form. It uses $x_m(t_{i-1})$ and the macroscopic state at the interface γ as input, evolves over the time-interval δt, and produces the state $x_m(t_i)$. The vector function $h(\bar{x}_s, x|_\gamma, dx/d\eta)$ represents the boundary conditions at the common interface between the macroscopic and microscopic domains, and $\bar{x}_s$ represents the stationary-state of the "coarse" realization, $\bar{x}$, of x_m. It is assumed that such stationary state exists and is independent of the initial microscopic state, *i.e.*, $x_m(t=0)$. The coarse variables, $\bar{x}$, can be projected onto the microscopic state variables, x_m, and vice versa, through the restriction and lifting operations $\bar{x} = L(x_m)$, $x_m = l(\bar{x})$, respectively (note that the lifting operation leads to a number of possible x_m for a given $\bar{x}$).

A general optimization problem for the multiscale system of Eqs. (4.44)–(4.47) can be formulated as (Varshney & Armaou, 2005a,b):

$$\min \mathcal{G}(x, \bar{x}_s, x, \delta\bar{t}_i) = \sum_i^n \int_\Omega G(x, \bar{x}_s, d, \delta\bar{t}_i)\,\mathrm{d}z$$

s.t.

$$\begin{aligned} &A(x) + f(x, d) = 0, \\ &g\left(x, \frac{\mathrm{d}x}{\mathrm{d}\eta}\right) = 0 \quad \text{on } \Gamma, \quad h\left(\bar{x}_s, x, \frac{\mathrm{d}x}{\mathrm{d}\eta}\right) = 0, \quad \text{on } \gamma \\ &p(x, d) \le 0, \quad \forall z \in \Omega_1 \end{aligned} \tag{4.48}$$

where $\mathcal{G}(x, \bar{x}_s, d, \delta\bar{t}_i)$ is the objective functional and measures the process performance at both macroscopic and microscopic levels and $p(x, \bar{x}, d)$ is the vector of inequality constraints which may include bounds on state and design variables. Time-intervals $\delta\bar{t}$ and design variables d_i (of Eq. (4.44)) are the optimization variables.

Finite-dimensional approximations to the semi-infinite dimensional program of Eq. (4.48) can be obtained through spatial discretization of the equality constraints and of the objective functional to formulate a finite-dimensional nonlinear program (NLP). Brute-force spatial discretization employing finite-difference/finite-element (FD/FE) techniques typically results in a large set of algebraic equations, and subsequent storage and computational requirements of the formulated NLP may become prohibitive requiring the use of specially designed algorithms for large-scale optimization problems. Inclusion of black-box timesteppers into the multiscale model further increases the computational demands. To address this issue, nonlinear order reduction for dissipative PDEs (Armaou & Christofides, 2002; Bendersky & Christofides, 2000) using Karhuenen–Loéve expansion (KLE) is coupled with *in situ* adaptive tabulation (Pope, 1997) to formulate reduced order multiscale models that can be employed to efficiently solve multiscale optimization problems.

The eigenspectrum of elliptic PDEs is characterized by a finite number of dominant eigenmodes (Armaou & Christofides, 2000; Baker & Christofides, 2000; Baker, Armaou, & Christofides, 2000; Christofides & Daoutidis, 1997), a property that constitutes the basis for computing dominant eigenmodes using KLE. These dominant eigenmodes can be identified empirically by applying KLE on an appropriate ensemble (for details about construction of the ensemble, the reader may refer to Armaou & Christofides, 2002; Bendersky & Christofides, 2000; Varshney & Armaou, 2006 and references therein) of PDE solution data. These eigenmodes, known as empirical eigenfunctions, can be subsequently employed as basis functions in the method of weighted residuals, to derive systems of algebraic equations, which have significantly smaller dimension than those derived using FD/FE discretization methods. Subsequently, the NLP can be solved using standard gradient-based or direct search algorithms.

The calculation of coarse stationary states, $\bar{x}_s$, through black-box timesteppers is usually a computationally expensive task. To facilitate efficient incorporation of black-box simulators, stationary-state coarse solution data (*i.e.*, $\bar{x}_s$) of black-box timesteppers are tabulated offline for the entire realizable region spanned by $x|_\gamma$. Necessary information from tabulated data can be obtained through interpolation, as required by the macroscopic solver. We employ adaptive tabulation to tabulate only the accessed region, which may be unknown *a priori*. The table is constructed *on demand*, when interpolation results of previously tabulated data are not accurate. The efficiency of adaptive tabulation increases if the accessed region is a small subset of the realizable region and contains domains with relatively large gradients.

An iterative solution algorithm that is applicable to a broad class of multiscale processes modeled by Eqs. (4.44)–(4.47) is outlined below (see Varshney & Armaou, 2005a,b for more details).

(1) Select an arbitrary (but physically consistent) initial condition $x_m(t=0)$ and $x|_\gamma$, and evolve the black-box timestepper till $\bar{x}$ reaches a stationary value (denoted as $\bar{x}_s$).
(2) Solve Eq. (4.44) subject to boundary conditions given by Eqs. (4.46) and (4.47), either analytically or numerically to obtain new $x|_\gamma$ denoted as x'_i.
(3) Repeat steps (1) and (2) to obtain x'_{i+1} until $x'_i - x'_{i+1}$ is below an acceptable tolerance.

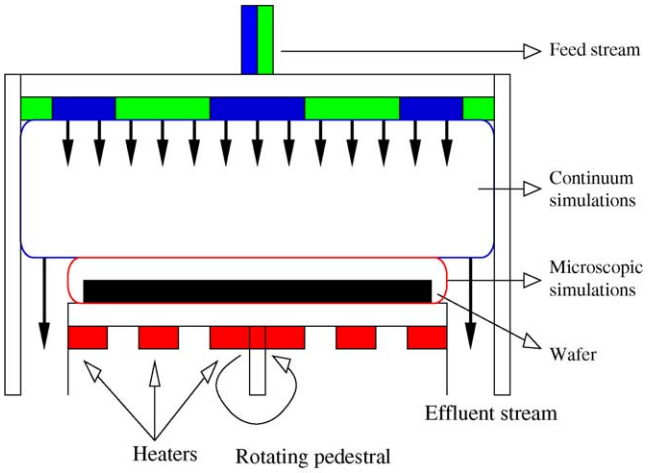

Fig. 13. Schematic of the reactor with split inlet configuration.

Subsequently, the reduced order model can be incorporated as an equality constraint into standard search algorithms such as BFGS, Luus–Jaakola, Hooke–Jeeves, *etc.*, to obtain the optimal solution.

4.2. Application to thin film growth

In this subsection, the above optimization methodology is applied to a conceptual thin-film growth process, where the objective is to compute an optimal time-varying process operation that simultaneously minimizes spatial thickness nonuniformity and surface roughness of the deposited film at the end of the process cycle; we note that this process is different from the one considered in Sections 1 and 2. Fig. 13 depicts the schematic of the reactor with split inlet configuration.

The bulk of the reactor is modeled using two-dimensional axisymmetric PDEs in cylindrical coordinates derived from continuum conservation principles. The surface of the growing film is modeled using kMC simulations. Fig. 13 also shows the domains of definition of the two models. It should be noted that the microscopic domain is infinitesimally thin. Substrate temperature profiles are manipulated using three circular heaters with heat being conducted in the in-between areas. Table 1 tabulates the reactor geometry and process conditions. Gaseous species A and B represent the precursors of a and b (components of compound semiconductor ab) respectively and are assumed to undergo the following gas phase reactions in the bulk of the reactor and gas-surface reactions on the wafer surface, shown in Table 2. Reaction $G1$ represents the thermal decomposition of precursor A into A' which adsorbs on the substrate (reaction $S1$). The rate-parameter for adsorption of A' (reaction $S1$) is assumed to be that of an ideal gas, *i.e.*, $k_a = s_0\sqrt{RT/2\pi M}$, where s_0 is the sticking coefficient. The rate of adsorption of B (reaction $S2$) is assumed to be equal to $S1$ so that the stoichiometry of the film is preserved. In addition to adsorption, diffusion and desorption of adsorbed species are other significant processes that affect the structure of the surface. The rate of desorption of surface species into the gas phase and the rate of surface diffusion are given by:

$$k_d^n = k_{d0}\, e^{-E_{d0}+n\Delta E/k_B T}, \quad k_m^n = \frac{k_B T}{h}\, e^{-E+n\Delta E/k_B T} \tag{4.49}$$

where h is the Planck's constant, E and E_{d0} the energy barriers for surface diffusion and desorption, respectively, ΔE the interaction energy between two neighboring adsorbed species and $n \in \{0, 1, 2, 3, 4\}$ is the number of nearest neighbors. The values of E, E_{d0}, ΔE and k_{d0} are taken as 2.5, 2.5, 0.5 eV and 1×10^{13}, respectively.

The macroscopic description of the process under consideration is given by the following conservation equations:

$$\begin{aligned}
&\nabla\cdot(\rho\mathbf{u}) = 0; \quad \nabla\cdot(\rho\mathbf{u}\,\mathbf{u}) - \nabla\cdot\mathbf{T} - \rho\mathbf{g} = 0\\
&\nabla\cdot(\rho\mathbf{u}T) = -\nabla\cdot\mathbf{q} - \sum_k h_k W_k \dot{w}\\
&\nabla\cdot(\rho\mathbf{u}Y_k) = -\nabla\cdot\mathbf{j}_k + W_k\dot{w}_k; \quad k \in \{1, 2, 3, 4\}\\
&\mathbf{j}_k = -D_k\rho\nabla Y_k - D_{T,k}\frac{\nabla T}{T}
\end{aligned} \tag{4.50}$$

where ρ is the gas phase density, $\mathbf{u}$ the fluid velocity vector, $\mathbf{T}$ the stress tensor, C_p the specific heat capacity, T the temperature, $\mathbf{q}$ the heat flux due to conduction and h_k, W_k and Y_k are the partial specific enthalpy, molecular weight and the mass fractions of gas species. $\dot{\omega}_k$ and $\mathbf{j}_k$ are the net production rate due to homogeneous reactions and mass flux respectively of species k. D_k and $D_{T,k}$ in the flux equation correspond to mass diffusion and thermal diffusion coefficients, respectively. The flux boundary condition at the deposition surface is given by:

$$\mathbf{j} = R_{ad} = k_a C_{A'}|_s - \langle k_d\rangle f(C_{a.s}, T, w_{A'A'}) \tag{4.51}$$

where R_{ad} is the net rate of adsorption, T_s the surface temperature, $\langle k_d\rangle$, $C_{A'}|_s$ and $C_{a.s}$ are the effective desorption rate, concentration of A' over the substrate and *average* surface concentration of adsorbed $a(s)$, respectively. Function f describes the influence of lateral interactions on the desorption rate, which cannot be ascertained without knowledge of surface structure. We employ kMC to account for the surface structure and esti-

Table 1
Design and process parameters

Reactor radius	2 in
Substrate radius (R_s)	1.5 in
Number of inlets	3
Substrate to inlet distance (z_0)	3 in
Reactor pressure	0.1 atm
Inlet & reactor wall temperature	300 K
Inlet velocity	80 cm/s
Substrate temperature (T_s)	900–1300 K
Inlet mass fraction of species A (X_A)	0.4×10^{-2}
Inlet mass fraction of species B (X_B)	0.6

Table 2
Process reaction scheme

Reaction	k_0	E
(G1) $A \rightarrow A' + C$	1×10^{14}	39.9
(S1) $A' \rightarrow a(s) + D$	–[a]	–
(S2) $B \rightarrow b(s)$	–[a]	–

[a] Rate calculated from kinetic theory of gases.

mate the right hand side of Eq. (4.51), which links the two levels of descriptions.

Initially, precursor A flows through the innermost inlet and B through the two outer inlets (ABB configuration of inlet). During the process operation, the two gas streams in the two innermost inlets can be interchanged to result in a distinct (BAB) inlet configuration. The transient evolution of the process following the switch is neglected (quasi-steady-state approximation). It is proposed that by optimally switching from ABB to BAB configuration and controlling the substrate temperature profile before and after the switching, both optimization objectives can be realized. Mathematically, the optimization problem can be formulated as:

$$\min F = \int_0^{R_0} \{w_1[\mathcal{T}(r) - \mathcal{T}_{\text{obj}}]^2 + w_2\mathcal{R}(r)\}\,\mathrm{d}r$$

s.t.

$$\mathcal{T} = \sum_{i=1}^{n} \delta\bar{t}_i R_{\text{dep}}; \quad \delta\bar{t}_i = \bar{t}_{i+1} - \bar{t}_i$$
$$u_k = \textstyle\sum_{i=1}^{n} u_{k,i}[H(\bar{t}_{i+1}) - H(\bar{t}_i)]; \quad k \in \{1, 2, 3\}$$
$$900 \le T_{\text{s}}(u_k) \le 1300; \quad R_{\text{dep}} = k_{\text{a}} C_{A'} \quad \text{at } \gamma \tag{4.52}$$

where F is the objective functional, T the thickness of film at the end of the process, T_{obj} the target thickness of the film, R_{dep} the deposition rate of species a, R the surface roughness of the deposited film, $\delta\bar{t}_i$ the time-interval for ith switching, T_{s} the surface temperature, u_k the magnitude of actuation and $H(\cdot)$ denotes the standard Heaviside function; the explicit dependence of T_{s} on u_k can be found in Varshney & Armaou (2005a). R_0 is the cutoff radius, which is taken to be a fraction of the substrate radius, thus discounting the unavoidable edge effects. The objective function penalizes any deviation of final film thickness from the target thickness (macroscopic objective) and high values of the spatially averaged roughness of the film (microscopic objectives). Additional constraints on the optimization problem arise from the reduced order process model, whose explicit form is omitted for brevity. The design variables of the optimization problem are the magnitudes of actuation u_k and the time-intervals $\delta\bar{t}_i$.

An ensemble of solution data ("snapshots") was generated by varying the substrate temperature (u_1, u_2 and u_3) for both ABB and BAB inlet configurations and solving the resulting system using the proposed multiscale algorithm. For the generation of snapshots, the macroscopic domain, Ω_1, was discretized using finite differences into 6201 nodes and the resulting system of nonlinear algebraic equations was solved using a Newton–Krylov-based solver. Specifically, an ensemble of 729×2 snapshots was generated. Three, 62 and 52 eigenfunctions were identified using KLE, respectively, for temperature and mass fraction profiles of A and A' across the reactor, which captured more than 99.999% of the energy of the ensemble. Hence, the reduced order model comprised of 117 (as opposed to 6201×3) nonlinear algebraic equations. Coarse data of kMC simulations was tabulated in accordance with *in situ* adaptive tabulation, as described earlier, which facilitated efficient linking. In order to account for the effect of radial variation of substrate temperature and the concentration of precursor over the substrate on the film microstructure, independent interpolation was performed at each macroscopic discretization node (a full order model would require independent kMC at each of these nodes, thereby significantly increasing the computational requirement). Depending upon the structure of the kMC simulator, the flow of information across the interface of the continuum and the discrete domains can be unidirectional or bidirectional. For the current process, numerical simulations established that inclusion of desorption into the kMC model had negligible effect on the macroscopic solution of the multiscale system (this one-way coupling characteristic of this process is in contrast to the fully coupled nature of the thin film growth process considered in Sections 1 and 2).

Hence, in the reduced order process, model desorption was not included. Under this assumption the flow of information was unidirectional and did not require multiple iterations. The resulting reduced optimization problem was solved using the Hooke–Jeeves search algorithm. The accuracy of the reduced order model was validated using the full order model under various operating conditions. However, direct comparison of full

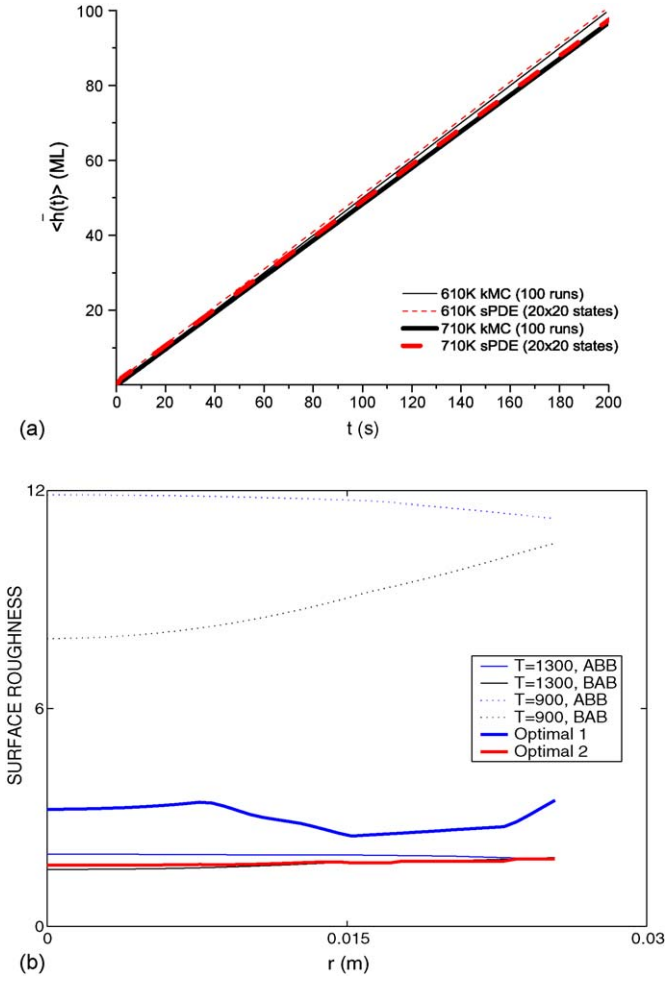

Fig. 14. Comparison of (a) deposition rate and (b) roughness profiles across the wafer surface, with macroscale (Optimal 1) and multiscale objective (Optimal 2).

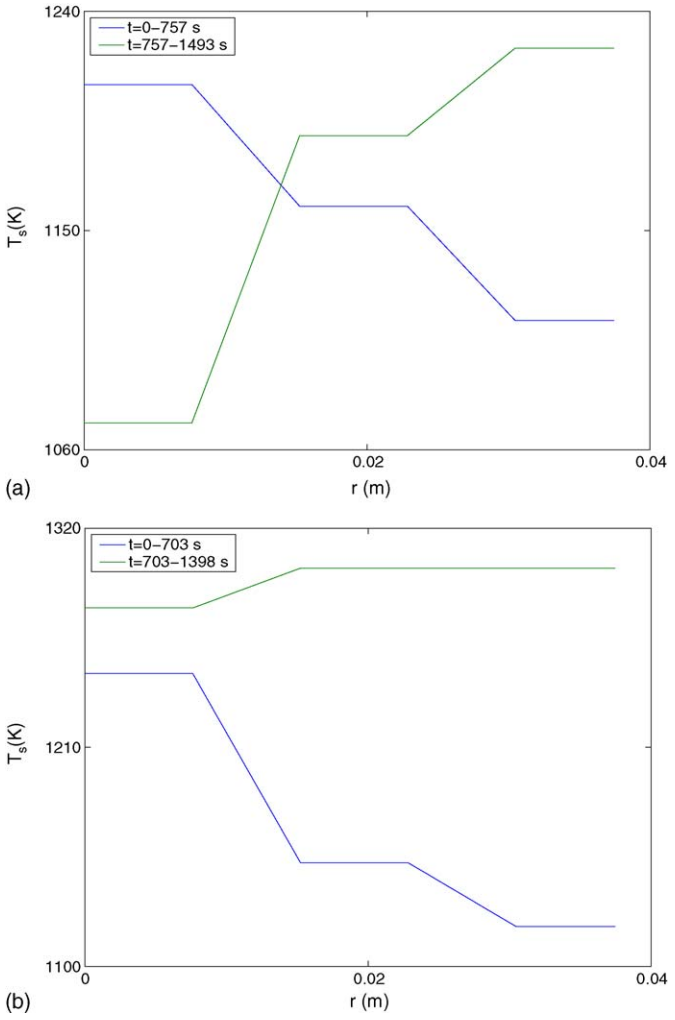

Fig. 15. Initial (blue line) and final (green line) optimal surface temperature profiles, for (a) macroscale only and (b) multiscale objective. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of the article.)

and reduced order formulations was impossible owing to the computational requirements of the full order model. Furthermore, once the empirical eigenfunctions and kMC-tables have been constructed, they can be repeatedly employed in multiple optimization problems.

The optimization problem of Eq. (4.52) was solved in two steps. Initially, spatial uniformity of the deposited film was the only optimization objective (*i.e.*, $\omega_1 \neq 0, \omega_2 = 0$). Subsequently, the microscopic objective was included into the optimization. The target film thickness T_{obj} was 5×10^{-6} m. Fig. 14a shows the final film thickness across the wafer surface obtained for the optimal process operation with macroscopic objective and multiscale objective (denoted as Optimal 1 and Optimal 2, respectively). For comparison purposes, the final film thickness-profile for time-invariant nominal process operation is also shown. Thickness nonuniformity, defined as $\sqrt{f(\mathcal{T}-\mathcal{T}_{obj})^2/\mathcal{T}_{obj}}$, for *ABB* inlet configuration was found to be 87.66 and 57.72% for substrate temperature 1300 and 900 K, respectively. The corresponding numbers were 215 and 188%, respectively, for *BAB* inlet configuration. For the optimal process operation, radial nonuniformity in the film was successfully reduced to 1.14 and 1.5% for the former and latter cases, respectively. The corresponding inlet switching times were 757 and 703 s, and the substrate temperature profiles before and after switching are shown in Fig. 15a and b, for macroscale and multiscale objectives, respectively.

Inclusion of the microscopic objective resulted in the overall increase of substrate temperature and the spatially averaged surface roughness of the film decreased from 3.5 (for Optimal 1) to 1.7 (for Optimal 2), shown in Fig. 14b. It should be noted that *BAB* configuration with $T_s = 1300$K would result in the film with the lowest surface roughness, however such operation is not optimal with respect to spatial film-thickness uniformity.

5. Conclusions and future research problems

In this manuscript, we have provided a tutorial overview of recently developed methods for control and optimization of complex process systems described by multiscale models. We primarily discussed methods developed in the context of our previous research work and used examples of thin film growth processes to motivate the development of these methods and illustrate their application. Specifically, we discussed: (a) a method for control of surface roughness in thin film growth using kinetic Monte-Carlo models, (b) a method for the construction of linear stochastic PDEs directly from data obtained by kinetic Monte-Carlo models and their subsequent use in the design of predictive controllers for roughness control in thin film growth and (c) a method for computationally efficient optimization of multiscale process systems with multiscale objectives based on reduction of the macroscopic PDEs using empirical eigenfunctions and reduction of the kinetic Monte-Carlo models via *in situ* adaptive tabulation—this optimization method was applied to a thin film growth process where the objective was to compute an optimal time-varying process operation that simultaneously minimizes spatial thickness nonuniformity and surface roughness of the deposited film at the end of the process cycle.

While significant progress has been made over the last 5 years on control and optimization of multiscale process systems, there are still many unresolved, challenging problems both in theory and applications. Here is a list of unresolved issues that, in our opinion, deserve to be studied:

- Construction of nonlinear stochastic differential equations from kinetic Monte-Carlo simulation data, as well as precise characterization of their accuracy, for processes where the molecular scale interactions give rise to highly nonlinear macroscopic behavior.
- Design of nonlinear/robust controllers based on stochastic linear/nonlinear models aiming at more direct and efficient compensation of model uncertainty and disturbances on operating conditions.
- Incorporation of dynamic behavior and uncertainty considerations in the multiscale optimization problem formulation and solution.

- Applications of the stochastic model construction, control and optimization methods to experimental thin film growth processes, as well as other complex processes where stochastic phenomena are important (*e.g.*, quantum-dot formation).

Acknowledgments

Financial support from the NSF (ITR), CTS-0325246, (PDC) and the Department of Chemical Engineering, Pennsylvania State University (AA), is gratefully acknowledged.

References

Armaou, A. (2005). Output feedback control of dissipative distributed processes via microscopic simulations. *Computers and Chemical Engineering*, *29*, 771–782.

Armaou, A., & Christofides, P. D. (1999). Plasma-enhanced chemical vapor deposition: Modeling and control. *Chemical Engineering Science*, *54*, 3305–3314.

Armaou, A., & Christofides, P. D. (2000). Wave suppression by nonlinear finite-dimensional control. *Chemical Engineering Science*, *55*, 2627–2640.

Armaou, A., & Christofides, P. D. (2002). Dynamic optimization of dissipative PDE systems using nonlinear order reduction. *Chemical Engineering Science*, *57*, 5083–5114.

Armaou, A., Kevrekidis, I. G., & Theodoropoulos, C. (2005). Equation-free controller designs for distributed complex/multiscale processes based on the gaptooth discretization scheme. *Computers and Chemical Engineering*, *29*, 731–740.

Armaou, A., Siettos, C. I., & Kevrekidis, I. G. (2004). Time-steppers and 'coarse' control of distributed microscopic processes. *International Journal of Robust and Nonlinear Control*, *14*, 89–111.

Baker, J., Armaou, A., & Christofides, P. D. (2000). Nonlinear control of incompressible fluid flow: Application to Burgers' equation and 2D channel flow. *Journal of Mathematical Analysis and Applications*, *252*, 230–255.

Baker, J., & Christofides, P. D. (2000). Finite dimensional approximation and control of nonlinear parabolic PDE systems. *International Journal of Control*, *73*, 439–456.

Bendersky, E., & Christofides, P. D. (2000). Optimization of transport-reaction processes using nonlinear model reduction. *Chemical Engineering Science*, *55*, 4349–4366.

Christofides, P. D. (2001). *Nonlinear and robust control of PDE systems: Methods and applications to transport-reaction processes*. Boston: Birkhauser.

Christofides, P. D. (2002a). *Model-based control of particulate processes*. The Netherlands: Kluwer Academic Publishers.

Christofides, P. D. (2002b). Control of distributed parameter systems [Special issue]. *Computers and Chemical Engineering*, *26*(7/8), 939–940.

Christofides, P. D., & Armaou, A. (2005). Control of multiscale and distributed process systems [Special issue]. *Computers and Chemical Engineering*, *29*(4), 687–688.

Christofides, P. D., & Daoutidis, P. (1997). Finite-dimensional control of parabolic PDE systems using approximate inertial manifolds. *Journal of Mathematical Analysis and Applications*, *216*, 398–420.

Drew, T. O., Webb, E. G., Ma, D. L., Alameda, J., Braatz, R. D., & Alkire, R. C. (2004). Coupled mesoscale–continuum simulations of copper electrodeposition in a trench. *AIChE Journal*, *50*, 226–240.

Edwards, S. F., & Wilkinson, D. R. (1982). The surface statistics of a granular aggregate. *Proceedings of the Royal Society of London, Series A*, *381*, 17–31.

Feller, W. (1975). *An introduction to probability theory and its applications*. New York: Wiley.

Fichthorn, K. A., & Weinberg, W. H. (1991). Theoretical foundations of dynamical Monte Carlo simulations. *Journal of Chemical Physics*, *95*, 1090–1096.

Gadgil, P. N. (1993). Single wafer processing in stagnation point flow CVD reactor: Prospects, constraints and reactor design. *Journal of Electronic Materials*, *22*, 171–177.

Gallivan, M. A., Goodwin, D. G., & Murray, R. M. (2004). Effective transition rates for epitaxial growth using fast modulation. *Physical Review B*, *70*(4), 045409.

Gallivan, M. A., & Murray, R. M. (2004). Reduction and identification methods for Markovian control systems, with application to thin film deposition. *International Journal of Robust and Nonlinear Control*, *14*, 113–132.

Gillespie, D. T. (1976). A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *Journal of Computational Physics*, *22*, 403–434.

Gillespie, D. T. (1992). A rigorous derivation of the chemical master equation. *Physica A*, *188*, 404–425.

Gilmer, G. H., & Bennema, P. (1972). Simulation of crystal growth with surface diffusion. *Journal of Applied Physics*, *43*, 1347–1360.

Lam, R., & Vlachos, D. G. (2001). Multiscale model for epitaxial growth of films: Growth mode transition. *Physical Review B*, *64*, 035401.

Lauritsen, K. B., Cuerno, R., & Makse, H. A. (2003). Noisy Kuramote–Sivashinsky equation for an erosion model. *Physical Review E*, *54*, 3577–3580.

Lou, Y., & Christofides, P. D. (2003a). Estimation and control of surface roughness in thin film growth using kinetic Monte-Carlo models. *Chemical Engineering Science*, *58*, 3115–3129.

Lou, Y., & Christofides, P. D. (2003b). Feedback control of growth rate and surface roughness in thin film growth. *AIChE Journal*, *49*, 2099–2113.

Lou, Y., & Christofides, P. D. (2004). Feedback control of surface roughness of GaAs (0 0 1) thin films using kinetic Monte-Carlo models. *Computers and Chemical Engineering*, *29*, 225–241.

Lou, Y., & Christofides, P. D. (2005a). Feedback control of surface roughness in sputtering processes using the stochastic Kuramoto–Sivashinsky equation. *Computers and Chemical Engineering*, *29*, 741–759.

Lou, Y., & Christofides, P. D. (2005b). Feedback control of surface roughness using stochastic PDEs. *AIChE Journal*, *51*, 345–352.

Marsili, M., Maritan, A., Toigo, F., & Banavar, J. R. (1996). Stochastic growth equations and reparametrization invariance. *Reviews of Modern Physics*, *68*, 963–983.

Mastny, E. A., Rawlings, J. B., & Kevrekidis, Y. G. (2005). Model based control methodologies for catalytic surface reactions. In *Proceedings of the American control conference* (pp. 2534–2539).

Melsa, J. L., & Sage, A. P. (1973). *An introduction to probability and stochastic processes*. Englewood Cliffs: Prentice-Hall.

Ni, D., & Christofides, P. D. (2005a). Construction of stochastic PDEs for feedback control of surface roughness in thin film growth. In *Proceedings of the American control conference* (pp. 2540–2547).

Ni, D., & Christofides, P. D. (2005b). Dynamics and control of thin film surface microstructure in a complex deposition process. *Chemical Engineering Science*, *60*, 1603–1617.

Ni, D., & Christofides, P. D. (2005c). Multivariable predictive control of thin film deposition using a stochastic PDE model. *Industrial and Engineering Chemistry Research*, *44*, 2416–2427.

Ni, D., & Christofides, P. D. (2006). Construction of stochastic PDEs and predictive control of surface roughness in thin film deposition. In *Model reduction and coarse-graining approaches for multiscale phenomena*. Berlin: Springer–Verlag, pp. 343–367.

Ni, D., Lou, Y., Christofides, P. D., Sha, L., Lao, S., & Chang, J. P. (2004). Real-time carbon content control for PECVD ZrO_2 thin film growth. *IEEE Transactions on Semiconductor Manufacturing*, *17*, 221–230.

Park, S., Kim, D., & Park, J. (2002). Derivation of continuum stochastic equations for discrete growth models. *Physical Review E*, *65*, 015102(R).

Pope, S. B. (1997). Computationally efficient implementation of combustion chemistry using *in situ* adaptive tabulation. *Combustion Theory Modelling*, *1*, 41–63.

Raimondeau, S., & Vlachos, D. G. (2000). Low-dimensional approximations of multiscale epitaxial growth models for microstructure control of materials. *Journal of Computational Physics*, *160*, 564–576.

Reese, J. S., Raimondeau, S., & Vlachos, D. G. (2001). Monte Carlo algorithms for complex surface reaction mechanisms: Efficiency and accuracy. *Journal of Computational Physics*, *173*, 302–321.

Rusli, E., Drews, T. O., Ma, D. L., Alkire, R. C., & Braatz, R. D. (2005). Robust nonlinear feedback–feedforward control of a coupled kinetic Monte-Carlo-finite difference simulation. In *Proceedings of the American control conference* (pp. 2548–2553).

Siettos, C. I., Armaou, A., Makeev, A. G., & Kevrekidis, I. G. (2003). Microscopic/stochastic timesteppers and "coarse" control: A KMC example. *AIChE Journal*, *49*, 1922–1926.

Siettos, C. I., Kazantzis, N., & Kevrekidis, Y. G. (2005). Equation-free, coarse-grained feedback linearization. In *Proceedings of the American control conference* (pp. 2554–2558).

Theodoropoulou, A., Adomaitis, R. A., & Zafiriou, E. (1999). Inverse model based real-time control for temperature uniformity of RTCVD. *IEEE Transactions on Semiconductor Manufacturing*, *12*, 87–101.

Varshney, A., & Armaou, A. (2005a). Multiscale optimization using hybrid PDE/kMC process systems with application to thin film growth. *Chemical Engineering Science*, *60*, 6780–6794.

Varshney, A., & Armaou, A. (2005b). Optimization of thin film growth using multiscale process systems. In *Proceedings of the American control conference* (pp. 2559–2565).

Varshney, A., & Armaou, A. (2006). Optimal operation of GaN thin-film epitaxy employing control vector parametrization. *AIChE Journal*, *52*, 1378–1391.

Vlachos, D. G. (1997). Multiscale integration hybrid algorithms for homogeneous–heterogeneous reactors. *AIChE Journal*, *43*, 3031–3041.

Vvedensky, D. D. (2003). Edwards–Wilkinson equation from lattice transition rules. *Physical Review E*, *67*, 025102(R).

Vvedensky, D. D., Zangwill, A., Luse, C. N., & Wilby, M. R. (1993). Stochastic equations of motion for epitaxial growth. *Physical Review E*, *48*, 852–862.

Wolfrum, P., Vargas, A., Gallivan, M., & Allgower, F. (2005). Complexity reduction of a thin film deposition model using a trajectory based nonlinear model reduction technique. In *Proceedings of the American control conference* (pp. 2566–2571).