2005 American Control Conference
June 8-10, 2005. Portland, OR, USA

Optimization of thin film growth using multiscale process systems

Amit Varshney and Antonios Armaou†
Department of Chemical Engineering
Pennsylvania State University, University Park, PA 16802

Abstract— **In this work we consider optimization problems for transport-reaction processes, when the cost function and/or equality constraints necessitate the consideration of phenomena that occur over widely disparate length scales. Initially, we develop multiscale process models that link continuum conservation laws with microscopic scale simulators. Subsequently, we combine nonlinear order reduction techniques for dissipative partial-differential equation systems with adaptive tabulation methods for microscopic simulators to reduce the computational requirements of the process description. The optimization problem is subsequently solved using standard search algorithms. The proposed method is applied to a representative thin film deposition process, where we compute optimal surface temperature profiles that simultaneously maximize deposition-rate uniformity (macroscale objective) and minimize surface roughness (microscale objective) across the film surface.**

I. Introduction

Numerous processes of industrial relevance require product quality specifications that are characterized by phenomena that evolve on disparate length scales. In general, a single mathematical description may not be applicable over the entire span of admissible length scales and in many cases appropriate closed-form descriptions may not be available for the whole domain. There may be regions inside the process domain where the only available descriptions are in the form of "black-box" timesteppers. These timesteppers may represent microscopic/mesoscopic evolution rules such as kinetic Monte-Carlo (kMC), Lattice-Botzmann (LB) or Molecular Dynamics (MD), or they may represent proprietary computational packages which interact only through input and output and whose inner details are unknown. Multiscale process models, that seamlessly combine available (closed form or black-box) descriptions at each length scale have to be developed when interplay between phenomena that occur at widely separated length scales need to be accounted for. Such models have been developed for fluid flows [11], [18], [22], [9]), moving contact line [14], transient fluid flows with heat transfer [7], crack propagation [5] and chemical vapor deposition [15], [28] to name a few. In most of these multiscale frameworks the inner (microscopic) and the outer (macroscopic) models evolve concurrently and interact through exchange of particles such that mass, momentum and energy remain conserved.

†Corresponding author armaou@psu.edu

Lately, considerable emphasis has been given to employ detailed process models for optimization and controller design. These process models are derived from dynamic conservation equations and usually take the form of parabolic/elliptic partial differential equations (PDEs) and constitute the equality constraints for subsequent process optimization analysis. These PDE equality constraints cannot be directly implemented, while employing standard optimization algorithms, and have to be discretized in space (and possibly time) to end up with algebraic equality constraints. Standard methods of discretization, such as finite differences/elements, can be utilized followed by sparse linear algebra based techniques to solve the resulting nonlinear program (NLP) [27]. However, one disadvantage of this method is that the NLPs thus formed are large in size, if accurate spatiotemporal behavior of the underlying PDEs has to be captured, and then their computational and storage requirements may be challenging. To overcome this difficulty, nonlinear order reduction strategies [3], [4], [1], [2] have been employed for the case of spatially distributed elliptic and parabolic PDEs, where spatial discretization takes advantage of the (few) dominant spatial patterns that appear in the solution of these PDEs [25]. These dominant (slow evolving and possibly unstable) eigenmodes, either obtained analytically (if possible) or empirically (e.g. through Karhuenen-Loéve expansion), can be utilized as basis functions during discretization using the method of weighted residuals. The NLPs, thus formulated, are significantly smaller in size. When such closed-form descriptions are unavailable, optimization methodologies using "black-box" functions have also been developed [17], [21]. The reader may also refer to [19], [20], where kMC methods were utilized for feedback control of surface roughness of thin films, and [10] where a model reduction approach for kMC simulations is applied to control surface roughness during thin film deposition. However, efficient optimization frameworks with multiscale process equality constraints and cost functionals are lacking.

Such methods will be advantageous to many processes including thin film deposition. A common issue with vapor phase epitaxy processes is that under nominal operating conditions the transport and reaction limitations severely affect the film deposition rate uniformity across the substrate. The quality of the film, determined by its electrical and magnetic properties, is also dependent on its crystal

0-7803-9098-9/05/$25.00 ©2005 AACC

structure[1]. Mathematical models for these quality control measures should span from macroscopic length scales (to account for reactor scale phenomena) to microscopic length scales (to account for phenomena that occur on the surface of the deposited film) [13].

In this paper we address the issue of optimization of processes where optimization objectives are separated by several orders of magnitude in length scale and hence necessitate multiscale process models. We focus on formulation of *consistent* multiscale and more importantly reduced-order-multiscale models, which conserve physical quantities across the interface of macroscopic and microscopic domains, in order to efficiently deal with the optimization objectives. We consider a conceptual vapor phase epitaxy process and calculate optimal substrate temperature profiles such that the grown thin films have a high degree of spatial uniformity and, simultaneously, low surface roughness. The macroscopic process description is obtained from steady-state conservation equations which assume the form of elliptic PDEs. The effect of possible parameter variations on the microstructure of the film is ascertained through mesoscale, kMC based, simulators that describe their evolution and have been coupled with PDE based macroscale models.

The paper is structured as follows. We begin with the description of a general multiscale process and the formulation of the optimization problem. We then proceed with a brief description of nonlinear order reduction of the macroscopic model and efficient tabulation of kMC solution data followed by a general algorithm to formulate a hybrid continuum/discrete multiscale model for the solution of the optimization problem. Subsequently, we apply the proposed methodology to a conceptual chemical vapor deposition process and optimize the substrate temperature profile to simultaneously maximize deposition rate uniformity and minimize roughness of the film.

II. Problem formulation

We consider a process for which distinct description at two different length scales is required. At the macroscopic level, the steady-state process description is given by spatially-distributed, elliptic PDEs and at the microscopic level, the time evolution of the process is given by a black-box function. We assume that the domains of definition of the macroscopic and the microscopic model, Ω_1 and Ω_2 respectively, do not overlap and share a common interface γ. Mathematically, the process is described as:

$$0 = \mathcal{A}(x) + f(x,d), \quad on\ \Omega_1 \tag{1}$$

$$\begin{gathered} x_m(t_i) = \Pi(x_m(t_{i-1}), \delta t, x|_\gamma), \quad on\ \Omega_2 \\ \delta t = t_i - t_{i-1} \end{gathered} \tag{2}$$

$$g(x, \frac{dx}{d\eta}) = 0, \quad on\ \Gamma \setminus \gamma \tag{3}$$

$$h(\bar{x}_s, x, \frac{dx}{d\eta}) = 0, \quad on\ \gamma \tag{4}$$

where $x(z) \in \mathbb{R}^n$ 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. $\mathcal{A}(x)$ 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, $g(x, \frac{dx}{d\eta})$ is a nonlinear vector function which is assumed to be sufficiently smooth, and η is the spatial direction perpendicular to the boundary Γ.

The black-box function, Π, describes the evolution of the microscopic process in the time interval $[t_{i-1}, t_i]$ which, using the microscopic state at time t_{i-i}, $x_m(t_{i-1})$ and the macroscopic state at the interface γ, $x|_\gamma$, provides the final state $x_m(t_i)$. Here it must be emphasized that this function interacts only through an input/output structure and it may be unknown in closed-form. The vector function $h(\bar{x}_s, x, \frac{dx}{d\eta})$ represents the boundary conditions at the common interface between the macroscopic and microscopic domains, γ, 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. $x_m(t=0)$.

Coarse variables, $\overline{x}$, are defined through the following restriction operator:

$$\overline{x} = L(x_m) \tag{5}$$

The inverse of restriction, termed as lifting, is defined by the following non-unique operator:

$$x_m = l(\bar{x}) \tag{6}$$

Note that the lifting operation leads to a number of possible x_m for a given $\bar{x}$. For a detailed analysis the reader may refer to [12]. A general optimization problem for the multiscale system described by Eqs.1-4 can be formulated as follows:

$$\begin{aligned} & min \int_{\Omega_1 \cup \Omega_2} G(x, \bar{x}_s, d) dz \\ s.t \quad & \mathcal{A}(x) + f(x,d) = 0, \\ & g(x, \frac{dx}{d\eta}) = 0 \ \ on\ \Gamma \\ & h(\bar{x}_s, x, \frac{dx}{d\eta}) = 0, \quad on\ \gamma \\ & p(x, \bar{x}_s, d) \leq 0, \quad \forall\ z \in \Omega_1 \end{aligned} \tag{7}$$

where $\int_{\Omega_1 \cup \Omega_2} G(x, \bar{x}_s, d) dz$ is the objective functional and measures the process performance at both macroscopic and microscopic levels and $p(x, \bar{x}_s, d)$ is the vector of inequality constraints which may include bounds on state and design variables.

[1]Here we consider surface roughness of the film as a representative measure of the defects in its microstructure which, in turn, affect its quality. However, roughness is only a crude way of estimating crystal defects.

III. Solution of Optimization Problem

Solution of the infinite dimensional program of Eq. 7, cannot be obtained directly through standard search algorithms. Equality and inequality constraints can be discretized in space to generate an approximate finite dimensional Nonlinear Program (NLP), which can be solved using standard gradient based or direct search algorithms. Brute-force spatial discretization of constraints employing finite differences or finite elements result in a large set of algebraic equations, and subsequent storage and computational requirements of NLP may become prohibitive requiring the use of large-scale optimization algorithms. Multiscale models that couple expensive black-box simulators further increase the computational demand. To address this issue, we couple nonlinear order reduction techniques for PDEs [3], [4] with *in situ* adaptive tabulation [23] to formulate reduced-order-multiscale models that can be employed to efficiently solve multiscale optimization problems.

A. Nonlinear order reduction of elliptic PDEs

A vast majority of the transport-reaction processes are described at macroscopic level by elliptic or parabolic PDEs. One important feature of elliptic PDEs is that their eigenspectrum is characterized by a finite number of dominant eigenmodes [6], [25]. These dominant eigenmodes can be identified empirically using Karhuenen-Loéve expansion on an appropriate ensemble (for details about construction of ensemble the reader may refer to [1], [2] and references therein) of PDE solution data. These eigenmodes, known as empirical eigenfunctions, can be subsequently employed as basis functions with the method of weighted residuals, to derive systems of algebraic equations, which are significantly lower in dimension than those derived using either finite differences or finite elements.

B. Adaptive tabulation of microscopic simulations

Usually calculation of coarse stationary states, $\bar{x}_s$, through black-box timesteppers is a computationally expensive task. In order to facilitate efficient incorporation of black-box simulators, stationary-state coarse solution data (i.e. $\bar{x}_s$) of black-box timestepper can be tabulated offline for the entire realizable region spanned by $x|_\gamma$. Necessary interpolation between the tabulated data can be done, as required by the macroscopic solver. To further reduce the computational requirements for the solution of the optimization problem, we employ adaptive tabulation and tabulate only the accessed region, which may not be known *a priori*. The table is constructed *on demand*, if interpolation based on previously tabulated data cannot be done accurately. 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. The details of the algorithm are omitted for brevity; interested author may refer [23]

C. Multiscale Solution Algorithm

In this section, we outline a solution algorithm that is applicable to a broad class of multiscale processes modeled by Eqs.1-4. Because Eq.4 involves variables from two distinct models, we formulate the following iterative procedure:

1) Start with an arbitrary (but physically consistent) initial condition $x_m(t = 0)$ and $x|_\gamma$, and evolve the black-box timestepper till $\overline{x}$ reaches a stationary value (denoted as $\bar{x}_s$). (We have already assumed that such a stationary state exists and is independent of $x_m(t = 0)$).
2) Solve Eq. 1 subject to boundary conditions given by Eqs. 3 and 4, 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 $\mid x'_i - x'_{i+1} \mid$ is below an acceptable tolerance.

Subsequently, the reduced order model can be incorporated into standard search algorithms such as successive quadratic programming (SQP), Luus-Jaakola, BFGS or Hooke-Jeeves to obtain the optimal solution [16].

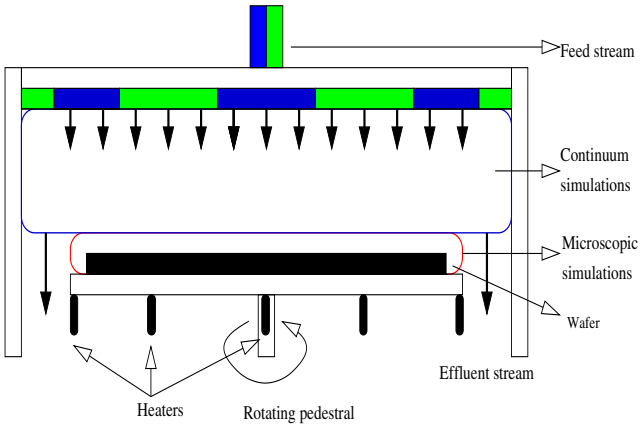

Fig. 1. Schematic of the reactor with split inlet configuration.

IV. Application to Thin Film Growth

The proposed optimization methodology is applied to a conceptual thin-film growth process, where we search for optimal substrate temperature profiles that simultaneously minimize deposition rate nonuniformity and surface roughness of the deposited film. Figure 1 depicts the schematic of the reactor and the domains of macroscopic and microscopic models. The bulk of the reactor can be accurately modeled using PDEs derived from continuum conservation principles, however, molecular motions at the surface of the substrate cannot be resolved using continuum PDEs and are thus modeled using kMC simulations. It should be noted that the microscopic domain is infinitesimally thin. Substrate temperature profiles can be manipulated using three circular heaters as shown in the figure. Assuming the specific reactor to be axisymmetric, macroscale descriptions

involving 2-dimensional elliptic PDEs in axial and radial directions are sufficient. Table I tabulates the reactor geometry and process conditions.

TABLE I

PROCESS CONDITIONS AND REACTOR GEOMETRY.

Reactor radius		2 in
Substrate radius	(R_s)	1.5 in
Number of inlets		2
inner inlet outer radius		0.5 in
outer inlet outer radius		1.5 in
Substrate to inlet distance	(z_0)	3 in
Reactor Pressure		0.1 atm
Reactor wall Temperature		300 K
Inlet Temperature		300 K
Inlet velocity		80 cm/s
Substrate temperature (T_s)		900-1300K
$X_A{}^*$		0.4×10^{-2}
$X_B{}^*$		0.6

*Inlet mass fractions of reactants

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 II.

TABLE II

GAS PHASE AND GAS-SURFACE REACTIONS

Reaction	k0	E
(G1) $A \rightarrow A' + C$	1×10^{14}	39.9
(S1) $A' \rightarrow a(s) + D$		0
(S2) $B \rightarrow b(s)$		

Reaction $G1$ represents the thermal decomposition of precursor A into A' which adsorbs on the substrate (reaction $S1$). Rate of surface reaction $S1$ (or the rate of adsorption of A') is assumed to follow that of an ideal gas with sticking coefficient s_0

$$k_a = s_0\sqrt{\frac{RT}{2\pi M}} \tag{8}$$

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 is calculated as:

$$\begin{aligned} k_d^n &= k_{d0}e^{-\frac{E_{d0}+n\Delta E}{k_BT}} \\ k_m^n &= \frac{k_BT}{h}e^{-\frac{E+n\Delta E}{k_BT}} \end{aligned} \tag{9}$$

where h is Planck's constant, E and E_{d0} are the energy barriers for surface diffusion and desorption respectively, ΔE is 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 eV, 2.5 eV, 0.5 eV and 1×10^{13}, respectively.

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

$$\begin{aligned} &\bigtriangledown\cdot(\rho\mathbf{u}) = 0; \quad \bigtriangledown\cdot(\rho\mathbf{u}\ \mathbf{u}) - \bigtriangledown\cdot\mathbf{T} - \rho\mathbf{g} = 0 \\ &\bigtriangledown\cdot(\rho\mathbf{u}T)] = -\bigtriangledown\cdot\mathbf{q} - \sum_k h_kW_k\dot{\omega}_k \\ &\bigtriangledown\cdot(\rho\mathbf{u}Y_k) = -\bigtriangledown\cdot\mathbf{j}_k + W_k\dot{\omega}_k, \\ &\qquad\qquad k \in \{1,2,3,4\} \\ &\mathbf{j}_k = -D_k\rho\bigtriangledown Y_k - D_{T,k}\frac{\bigtriangledown T}{T} \end{aligned} \tag{10}$$

where ρ is the gas-phase density, $\mathbf{u}$ is the fluid velocity vector, $\mathbf{T}$ is the stress tensor, C_p is the specific heat capacity, T is the temperature, $\mathbf{q}$ is 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 above conservation equations along with the associated boundary conditions, complete the macroscopic model. The flux boundary condition at the deposition surface is given by:

$$\mathbf{j} = R_{ad} \tag{11}$$

where R_{ad} is the net rate of adsorption and surface temperature, T_s, is the design variable. We assume the rate of surface reaction $S1$ to be independent of surface configuration and express R_{ad} as [28]

$$R_{ad} = k_aC_{A'}|_s - < k_d > f(C_{a.s}, T, w_{A'A'}) \tag{12}$$

where $< k_d >$, $C_{A'}|_s$ 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 estimate the right hand of Eq. 11, which links the two levels of descriptions.

kMC approximates the solution of the stochastic master equation [8] through Monte-Carlo sampling

$$\frac{\partial P(\sigma,t)}{\partial t} = \sum_{\sigma'} W(\sigma' \rightarrow \sigma)P(\sigma',t) - \sum_{\sigma'} W(\sigma \rightarrow \sigma')P(\sigma,t)$$

where σ and σ' are system configurations and $P(\sigma,t)$ is the probability that the system is in state σ at time t, and $W(\sigma \rightarrow \sigma')$ is the probability per unit time of transition from σ to σ'. It is assumed that at any instant, only a single event (out of all possible events) occurs, according to its relative probability. After each event, time is incremented by δt, given as:

$$\delta t = -\frac{lnr}{\sum_i \Psi_i} \tag{13}$$

where r is a random number between 0 and 1 and Ψ_i is the propensity function of event i. The summation in the denominator is carried over all possible events and transition probabilities are adjusted after each event. For the present case, Eq.13 becomes

$$\delta t = -\frac{lnr}{k_a N_T + \sum_{n=0}^{4} k_{m,n} N_n} \tag{14}$$

where N_T is the total number of surface sites and N_n is the number of sites with n nearest neighbors. The roughness of the film is computed from [24], [19]:

$$\begin{aligned}\mathcal{R} = \tfrac{1}{2N^2}\textstyle\sum_{i,j}(|h_{i+1,j} - h_{i,j}| + |h_{i-1,j} - h_{i,j}|) + \\ \tfrac{1}{2N^2}\textstyle\sum_{i,j}(|h_{i,j+1} - h_{i,j}| + |h_{i,j-1} - h_{i,j}|)\end{aligned} \tag{15}$$

where $h_{i,j}$ is the number of atoms adsorbed at the $(i,j)^{th}$ surface site.

Uniformity in the rate of deposition of species a and minimization of the surface roughness of the film are the two important process objectives during thin film deposition. We investigate the effect of substrate temperature on them and search for the optimal temperature profile that simultaneously achieves the two objectives. Mathematically, the optimization problem is formulated as:

$$\begin{aligned} min \;\; F =& w_1 \int_0^{R_0} \left\{R_{dep}(r) - \bar{R}_{dep}\right\}^2 dr \\ &+ w_2 \int_0^{R_0} \mathcal{R}(r) dr + w_3 \int_0^{R_0} (\mathcal{R}(r) - \bar{\mathcal{R}})^2 dr \\ & \text{s.t.} \\ & R_{dep} = k_a C_{A'} \;\; at \;\; \gamma \\ & \bar{R}_{dep} = \frac{1}{R_0}\int_0^{R_0} R_{dep}(r) dr \\ & 900K \leq T_s \leq 1300K \end{aligned} \tag{16}$$

where F is the objective functional, R_{dep} is the deposition rate of the thin film, $\mathcal{R}$ is the surface roughness of the deposited film and T_s is the surface temperature. 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 non-uniformities in the deposition rate (macroscopic objective) and the surface roughness across the substrate, as well as 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 variable for the above optimization problem, T_s, was assumed to take the following form:

$$T_s = u_1 \delta_d(0) + u_2 \delta_d(r - R_0/2) + u_3 \delta(r - R_0) \tag{17}$$

where u_1, u_2 and u_3 are the magnitudes of actuation and δ_d is the Dirac delta function.

Remark 1: In practice, depending upon the specific process under consideration, additional constraints may be considered. For example, temperature gradient constraints over the wafer surface may be imposed to prevent the film degradation due to nonuniform thermal expansion. Such constraints can be readily incorporated into the optimization problem formulation outlined above.

A. Reduced order model

In this section we first describe the coupling scheme of the macroscopic PDEs with the kMC simulator for the solution of the multiscale process model. As mentioned earlier, the link is provided by the right hand side of Eq. 11. The general algorithm employed for steady-state multiscale calculations, is presented below:

1) Solve the macroscopic model, using maximum flux boundary condition at the surface (i.e. Eq. 12 with $< k_d >= 0$) to obtain an initial estimate of concentration of adsorbing species over the surface.
2) Calculate the steady-state net surface flux by running kMC simulations for sufficiently long time.
3) Solve the macroscopic model using the flux calculated in step 2 as RHS of the Eq. 11, and obtain modified concentration profiles of all species over the surface.
4) Using the concentrations obtained in step 3, repeat steps 2 and 3 till the concentrations used in step 2 and the concentration obtained in step 3 are within the acceptable tolerance.

An ensemble of solution data ("snapshots") was generated by varying the substrate temperature (through u_1, u_2 and u_3) and solving the resulting system according to the algorithm presented above. For the generation of snapshots, the macroscopic domain, $\Gamma 1$, 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 snapshots was generated. Karhuenen-Loéve expansion identified 3, 25 and 40 eigenfunctions, 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. The reduced order model, hence, comprised of 68 (as opposed to 6201×3) nonlinear algebraic equations. Coarse data of kMC simulations was tabulated employing $in\ situ$ adaptive tabulation, described earlier, which facilitated efficient linking of the two scales (from a computational perspective). The resulting reduced optimization problem was solved using Hooke-Jeeves search algorithm.

Remark 2: 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 case, numerical simulations established that inclusion of desorption into the kMC model had negligible effect on the macroscopic solution of the multiscale system. 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. For further details on the proposed method, refer to [26].

V. Results

Choice of the size of lattice for kMC involves tradeoff (similar to the choice of grid size in case of finite elements), since large lattice sizes increase the computational demand. Conversely, kMC with small lattice size has significant stochastic uncertainty. Figure 2, compares the time evolution of surface roughness computed using 50×50, 75×75 and 100×100 lattices. We observe that the results from all the three lattice sizes are comparable. Hence we chose 75×75 lattice for the rest of the computations. The pertinent simulation parameters are listed in the figure caption.

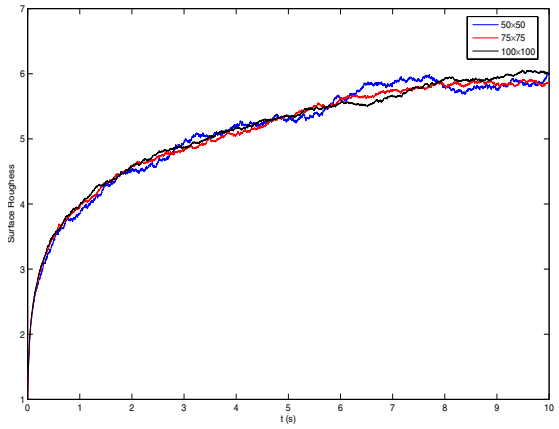

Fig. 2. Surface roughness as a function of time for various lattice sizes (T_s=1100K; adsorption rate=10 atoms/site-sec).

In Figure 3, surface deposition rate profiles of species a are presented for $T_s = 900K$ and $T_s = 1300K$ and compared against that of optimal surface temperature profiles obtained from the solution of the problem of Eq.16 by accounting first only for the macroscopic objective (i.e., $w_2 = w_3 = 0$), and secondly for the combined macroscopic-microscopic objective (i.e., $w_1, w_2 \neq 0, w_3 = 0$). Deposition rate non-uniformity is defined as:

$$[R_{dep}(r=0) - min(R_{dep}(r))]/R_{dep}(r=0)$$

and is computed to be 37.86%, 26.05%, 6.79% and 13.21% for $T_s = 1300K$, $T_s = 900K$ and the two optimal cases, respectively. The corresponding optimal temperature and surface roughness profiles are shown in Figures 4 and 5 respectively. For constant substrate temperature operation, the concentration of the adsorbed species and hence the deposition rate reduces across the substrate radius, except near the edge where due to high convective mass transfer the deposition rate increases sharply. This unavoidable "edge" effect is not presented in the figure. We observe that, temperature gradients near the substrate induced by radial variations in substrate temperature cause preferential diffusion of species radially, which lead to modified deposition rates. The solution of the optimization problem demonstrates that optimal variation of the substrate temperature (by controlling the magnitude of actuation) can lead to significant improvement in the uniformity of the grown film. Higher overall surface temperature is required if reduction of average surface roughness is a concurrent optimization objective (see Figure 4), since increased surface diffusion at high temperatures leads to smoothening of the deposited film. However, the modified temperature gradients (especially near the substrate) may not favor axial thermal diffusion to the desired extent and consequently deposition rate uniformity may be compromised. The optimal temperature profile with respect to the above concurrent objective reduces the average surface roughness from 9.26 (for macroscale-only case) to 6.59, increasing the nonuniformity in deposition rate from 6.79% to 13.21%. It should be noted that process operation with substrate temperature near $1300K$ results in notable reduction in surface roughness, however, it is not optimal with respect to the deposition rate uniformity.

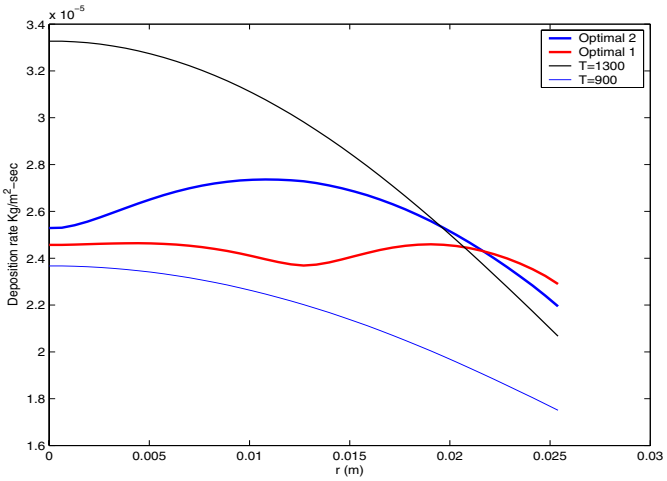

Fig. 3. Comparison of deposition rate profiles with macroscale (Optimal 1) and macroscale-microscale optimization objectives (Optimal 2).

VI. Conclusion

Nonlinear order reduction techniques for elliptic partial differential equations were employed together with adaptive tabulation method for microscopic/mesoscopic simulators to derive reduced-order continuum-discrete hybrid models for computationally efficient solution of optimization problems when the cost function and/or equality constraints necessitate the consideration of phenomena that occur over widely disparate length scales. The proposed methodology was applied to a conceptual thin film growth process to simultaneously maximize deposition rate uniformity and minimize average surface roughness by varying the temperature across the wafer substrate. A hybrid model was developed from conservation laws for heat, mass and momentum for the gas phase and kMC simulations were used to describe the film growth.

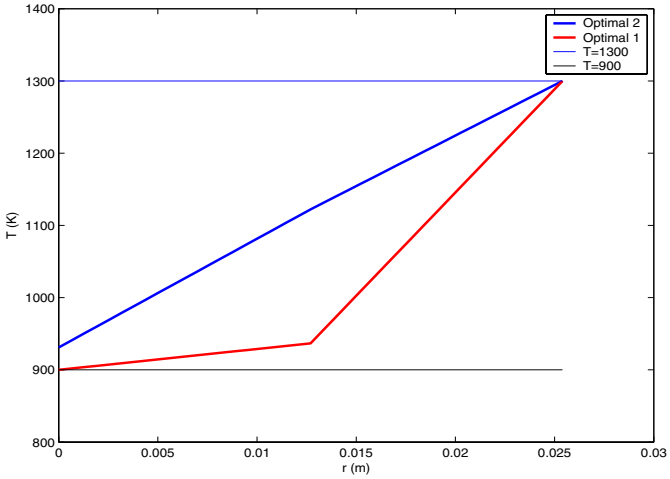

Fig. 4. Optimal substrate temperature profiles across the wafer surface.

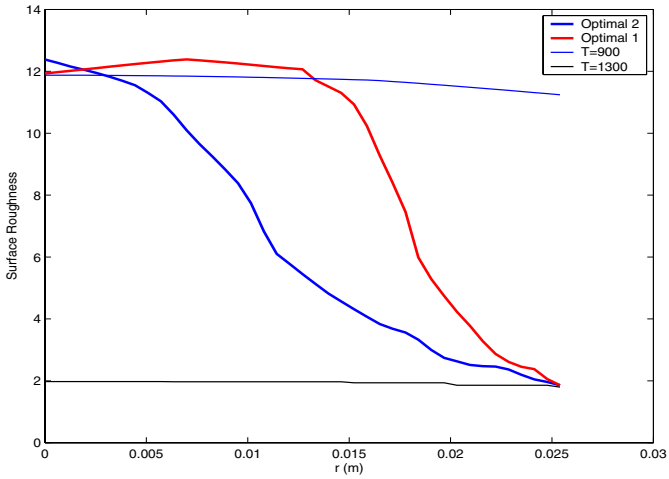

Fig. 5. Surface roughness profiles with macroscale (Optimal 1) and macroscale-microscale optimization objectives (Optimal 2).

VII. Acknowledgements

Financial support for this work from the Department of Chemical Engineering, Pennsylvania State University is gratefully acknowledged. We would also like to thank Professor Costas Maranas for providing us with the computational resources.

References

[1] A. Armaou and P. D. Christofides. Dynamic optimization of dissipative PDE systems using nonlinear order reduction. *Chem. Eng. Sci.*, 57:5083–5114, 2002.

[2] A. Armaou and A. Varshney. Dynamic optimization of dissipative PDEs using control vector parameterization: Application to GaN thin film epitaxy. In *Proceedings of the American Control Conference*, pages 279–286, Boston, Massachusetts, 2004.

[3] E. Bendersky and P.D. Christofides. A computationally efficient method for optimization of transport-reaction processes. *Comp. & Chem. Eng.*, 23(s):447–450, 1999.

[4] E. Bendersky and P.D. Christofides. Optimization of transport-reaction processes using nonlinear model reduction. *Chem. Eng. Sci.*, 55:4349–4366, 2000.

[5] J. Q. Broughton, F. F. Abraham, N. Bernstein, and E. Kaxiras. Concurrent coupling of length scales: Methodology and application. *Phys. Rev. B*, 60:2391–2403, 1999.

[6] P. D. Christofides. *Nonlinear and Robust Control of PDE Systems: Methods and Applications to Transport-Reaction Processes*. Birkhäuser, Boston, 2001.

[7] R. Delgado-Buscalioni and P. V. Coveney. Continuum-psrticle hybrid coupling for mass. momentum, and energy transfers in unsteady fluid flow. *Physical Review E*, 67:046704, 2003.

[8] K. A. Fichthorn and W. H. Weinberg. Theoretical foundations of dynamical Monte Carlo simulations. *J. Chem. Phys.*, 95:1090–1096, 1991.

[9] E. G. Flekkoy, G. Wagner, and J. Feder. Hybrid model for combined particle and continuum dynamics. *Europhys. Lett.*, 52:271–276, 2000.

[10] M. A. Gallivan. Optimization, estimation, and control of kinetic Monte Carlo simulations of thin film deposition. In *Proceedings of the 42nd IEEE Conference on Decision and Control*, pages 3437–3442, 2003.

[11] A. L. Garcia, J. B. Bell, W. Y. Crutchfield, and B. J. Alder. Adaptive mesh and algorithm refinement using direct simulation monte carlo. *J. Comp. Phys.*, 154:134–155, 1999.

[12] C. W. Gear, J. M. Hyman, P. G. Kevrekidis, Runborg. O., K. Theodoropoulos, and I. G. Kevrekidis. Equation-free multiscale computation: enabling microscopic simulators to perform system-level tasks. *Comm. Math. Sciences*, 2002.

[13] M. Gummalla, M. Tsapatsis, J. J. Watkins, and D. G. Vlachos. Multiscale hybrid modeling of film deposition within porous substrates. *AIChE J.*, 50:684–695, 2004.

[14] N. G. Hadjiconstantinou. Hybrid atomistic-continuum formulations and the moving contact line problem. *J. Comp. Phys.*, 154:245–265, 1999.

[15] K. F. Jensen, S. T. Rodgers, and R. Venkataramani. Multiscale modeling of thin film growth. *Cur. Opin. Sol. State & Mat. Sci.*, 3:562–569, 1998.

[16] C. T. Kelley. *Iterative Methods for Optimization*, volume 18 of *Frontiers in Applied Mathematics*. SIAM, Philadelphia, PA, USA, 1999.

[17] C. T. Kelley and E. W. Sachs. Truncated newton methods for optimization with inaccurate functions and gradients. *J. Opt. Theory Appl.*, 116:83–98, 2003.

[18] J. Li, D. Liao, and S. Yip. Nearly exact solution for coupled continuum/MD fluid simulation. *J. Comp. mat. Des.*, 6:95–102, 1999.

[19] Y. Lou and P. D. Christofides. Estimation and control of surface roughness in thin film growth using kinetic Monte-Carlo models. *Chem. Eng. Sci.*, 58:3115–3129, 2003.

[20] Y. Lou and P. D. Christofides. Feedback control of growth rate and surface roughness in thin film growth. *AIChE J.*, 49:2099–2113, 2003.

[21] C. A. Meyer, C. A. Floudas, and A. Neumaier. Global optimization with nonfactorable constraints. *Ind. Eng. Chem. Res.*, 41:6413–6424, 2002.

[22] S. T. O'Connell and P. A. Thompson. Molecular dynamic-continuum hybrid computations: A tool for studying complex fluid flows. *Phys. Rev. E*, 52:5792–5795, 1995.

[23] S. B. Pope. Computationally efficient implementation of combustion chemistry using $in\ situ$ adaptive tabulation. *Combust. Theory Modelling*, 1:41–63, 1997.

[24] S. Raimondeau and D. G. Vlachos. Low-dimensional approximations of multiscale epitaxial growth models for microstructure control of materials. *J. Comp. Phys.*, 160:564–576, 2000.

[25] R. Temam. *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*. Springer-Verlag, New York, 1988.

[26] A. Varshney and A. Armaou. Multiscale optimization using hybrid PDE/kMC process systems with application to thin film growth. *Chemical Engineering Science*, submitted, 2004.

[27] S. Vasantharajan, J. Viswanathan, and L. T. Biegler. Reduced successive quadratic programming implementation for large-scale optimization problems with smaller degrees of freedom. *Comp. Chem. Eng.*, 14:907–915, 1990.

[28] D. G. Vlachos. Multiscale integration hybrid algorithms for homogeneous-heterogeneous reactors. *AIChE J.*, 43:3031–3041, 1997.