OPTIMAL OPERATION OF THIN FILM GROWTH WITH MULTISCALE PROCESS OBJECTIVES

Amit Varshney and Antonios Armaou [1]

Department of Chemical Engineering, The Pennsylvania State University, University Park, PA-16802, USA

Abstract: The issue of optimal time-varying operation for transport-reaction processes is considered, when the cost functional and/or equality constraints necessitate the consideration of phenomena that occur over disparate length scales. Multiscale process models are initially developed, linking continuum conservation laws with microscopic scale simulators. Subsequently, order reduction techniques for dissipative partial-differential equations are combined with adaptive tabulation methods for microscopic simulators to reduce the computational requirements of the optimization problem, which is then solved using standard search algorithms. The method is demonstrated on a thin film deposition process, where optimal surface temperature profiles and inlet switching times that simultaneously maximize thickness uniformity and minimize surface roughness across the film surface are computed.

Keywords: optimization, distributed-parameter systems, order reduction, process control, complex systems

1. INTRODUCTION

Ever tightening product quality specifications have motivated the development of process models using increasingly complex and detailed process descriptions such as Molecular Dynamics (MD), Monte Carlo/kinetic Monte Carlo (MC/kMC) simulations, because of their ability to describe mesoscopic/microscopic length-scale phenomena which are inaccessible with traditional continuum laws. However, the computational requirements associated with these models limit their applicability only to micro-size domains. To this end, hybrid continuum/atomistic multiscale models have been developed that augment the continuum macroscopic description of a process by embedding microscopic or mesoscopic descriptions *only* at small domains where finer, than the one provided by macroscopic models, resolution is required (Garcia *et al.*, 1999; Li *et al.*, 1999; O'Connell and Thompson, 1995; Flekkoy and Feder, 2000; Broughton *et al.*, 1999; Vlachos, 1997).

Steady-state macroscopic process models are typically derived from conservation laws for mass, momentum and energy, and comprise of spatially-distributed partial differential equations (PDEs). On the other hand, microscopic models are usually unavailable in closed-form and their "integration" involves successively advancing the state of the system by small incremental time-steps. Note that MD, kMC and Lattice-Boltzmann-based "black-box" timesteppers fit into the above description. The standard optimization procedure for macroscopic (PDE-based) process descriptions is to discretize the PDEs using finite difference/elements (FD/FE) and subsequent solution of the large-scale nonlinear program (NLP) using techniques such as reduced gradient or reduced successive quadratic programming (Floudas and Pardalos, 1992). An alternate methodology is to formulate reduced-order approximate NLPs by spatial discretization combining the method of

[1] Corresponding author armaou@psu.edu

weighted residuals with empirical eigenfunctions and solve the NLP using standard search algorithms (Bendersky and Christofides., 2000). To address the issue of model unavailability for process when black-box simulations form part of the equality constraints, optimization methodologies for noisy black-box functions have also been developed (Kelley and Sachs, 2003; Meyer *et al.*, 2002). However, optimization methodologies for multiscale process models are still lacking.

The current work addresses the issue of efficient solution of optimization problems when cost functional and/or equality constraints span multiple length scales and necessitate multiscale process models. Order-reduction techniques for macroscopic process models are linked with adaptive tabulation schemes for the solution data from microscopic models to derive computationally efficient multiscale models which form the equality constraints of the optimization problem. The proposed method is applied to a conceptual thin-film epitaxy process with two simultaneous objectives that span multiple length scales: a) to maximize the thickness uniformity of the deposited film (macroscopic objective), and b) to minimize the surface roughness of the deposited film (microscopic objective) across the wafer surface at the end of the process cycle. Mathematically the process is modeled using continuum conservation laws and the microscopic film-surface processes are modeled using kMC simulations. Time-varying substrate temperature radial profile and inlet concentration profiles of the precursors are employed as design variables and quasi-steady state process operation is assumed.

2. PROBLEM FORMULATION

Mathematically, the spatially distributed and multiscale process can be represented as:

$$\begin{aligned} 0 &= \mathcal{A}(x) + f(x,d), \quad on\ \Omega_1 \\ d &= \sum_i^n d_i(z)(H(t-\bar{t}_i) - H(t-\bar{t}_{i+1})) \end{aligned} \tag{1}$$

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

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

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

Eq.1 and Eq.2 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. $\mathcal{A}(x)$ 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 and $\bar{t}_i$ is the time-instant when the design variables are varied. $g(x, \frac{dx}{d\eta})$, defined on the boundary $\Gamma \setminus \gamma$ 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.1 relax, t_i^s, is negligible in comparison to $\delta\bar{t}_i = \bar{t}_{i+1} - \bar{t}_i$, implying that the process is operating under quasi-steady-state conditions. Function, Π, can be thought of as a black-box timestepper, which interacts 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 state $x_m(t_i)$. The vector function $h(\bar{x}_s, x|_\gamma, \frac{dx}{d\eta})$ 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 $\bar{x} = L(x_m), \quad x_m = l(\bar{x})$, respectively (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 (Kevrekidis *et al.*, 2003)).

A general optimization problem for the multiscale system of Eqs.1-4 can be formulated as:

$$\begin{aligned} &min\ \mathcal{G}(x, \bar{x}_s, d, \delta\bar{t}_i) = \sum_i^n \int_\Omega G(x, \bar{x}_s, d, \delta\bar{t}_i) dz \\ &\qquad s.t \\ &\mathcal{A}(x) + f(x,d) = 0, \\ &g(x, \frac{dx}{d\eta}) = 0 \ \ on\ \Gamma, \quad h(\bar{x}_s, x, \frac{dx}{d\eta}) = 0, \quad on\ \gamma \\ &p(x,d) \le 0, \ \ \forall\ z \in \Omega_1 \end{aligned} \tag{5}$$

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 $\delta\bar{t}_i$ and design variables d_i (of Eq.1) are the optimization variables.

3. OPTIMIZATION PROBLEM SOLUTION

Finite-dimensional approximations to the semi-infinite dimensional program of Eq.5 can be obtained through spatial discretization of equality constraints and objective functional to formulate a Nonlinear Program (NLP). Brute-force spatial discretization employing FD/FE result 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 PDEs (Bendersky and 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 (Christofides, 2001), a property that constitutes the basis of KLE. These dominant eigenmodes can be identified empirically using KLE on an appropriate ensemble (for details about construction of ensemble the reader may refer to (Bendersky and Christofides., 2000; Armaou and Christofides, 2002) 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 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.1-4 is outlined below.

(1) Select 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 $\bar{x}_s$).
(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 $x'_i - x'_{i+1}$ is below an acceptable tolerance.

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. APPLICATION TO THIN FILM GROWTH

The proposed 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. Figure 1 depicts the schematic of the reactor with split inlet configuration. The bulk of the reactor is modeled using

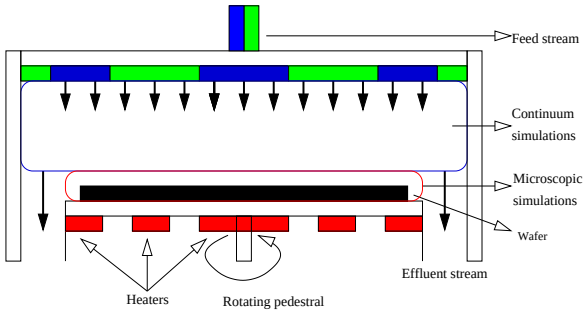

Fig. 1. Schematic of the reactor with split inlet configuration.

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)	900-1300K
Inlet mass fraction of species A	(X_A)	0.4×10^{-2}
Inlet mass fraction of species B	(X_B)	0.6

2-dimensional axisymmetric PDEs in cylindrical coordinates derived from continuum conservation principles. The surface of the growing film is modeled using kMC simulations. Figure 1 also shows the domains of description 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.

Table 2. Process reaction scheme

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 $S1$). The rate-parameter for adsorption of A' (reaction $S1$) is assumed to follow that of an ideal gas, i.e. $k_a = s_0\sqrt{\frac{RT}{2\pi M}}$, 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 is calculated as:

$$k_d^n = k_{d0}e^{-\frac{E_{d0}+n\Delta E}{k_B T}}, \quad k_m^n = \frac{k_B T}{h}e^{-\frac{E+n\Delta E}{k_B T}} \tag{6}$$

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:

$$\begin{aligned}
&\bigtriangledown \cdot (\rho \mathbf{u}) = 0; \quad \bigtriangledown \cdot (\rho \mathbf{u}\ \mathbf{u}) - \bigtriangledown \cdot \mathbf{T} - \rho \mathbf{g} = 0 \\
&\bigtriangledown \cdot (\rho \mathbf{u} T)] = - \bigtriangledown \cdot \mathbf{q} - \sum_k h_k W_k \dot{\omega} \\
&\bigtriangledown \cdot (\rho \mathbf{u} Y_k) = - \bigtriangledown \cdot \mathbf{j}_k + W_k \dot{\omega}_k; \quad k \in \{1, 2, 3, 4\} \\
&\mathbf{j}_k = -D_k \rho \bigtriangledown Y_k - D_{T,k} \frac{\bigtriangledown T}{T}
\end{aligned} \tag{7}$$

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 flux boundary condition at the deposition surface is given by (Vlachos, 1997):

$$\mathbf{j} = R_{ad} = k_a C_{A'}|_s - < k_d > f(C_{a.s}, T, w_{A'A'}) \tag{8}$$

where R_{ad} is the net rate of adsorption, T_s is the surface temperature, $< k_d >$, $C_{A'}|_s$ 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 side of Eq. 8, which links the two levels of descriptions.

kMC approximates the solution of the stochastic master equation (Fichthorn and Weinberg, 1991) through Monte-Carlo sampling

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

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 $W(\sigma, \sigma')$ 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:

$$\delta t = -\frac{lnr}{\sum_i \Psi_i} = -\frac{lnr}{k_a N_T + \sum_{n=0}^{4} k_{m,n} N_n} \tag{9}$$

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. N_T is the total number of surface sites and N_n is the number of sites with n nearest neighbors. The surface roughness is computed from (Lou and Christofides, 2003):

$$\begin{aligned}
\mathcal{R} = &\frac{1}{2N_T} \sum_{i,j} (|h_{i+1,j} - h_{i,j}| + |h_{i,j-1} - h_{i,j}| \\
&+ |h_{i-1,1} - h_{i,j}| + |h_{i,j+1} - h_{i,j}|)
\end{aligned} \tag{10}$$

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

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:

$$
\begin{aligned}
&min \;\; F = \int_0^{R_0} \left\{ w_1 [\mathcal{T}(r) - \mathcal{T}_{obj}]^2 + w_2 \mathcal{R}(r) \right\} dr \\
&\qquad\qquad \text{s.t.} \\
&\mathcal{T} = \sum_{i=1}^{n} \delta\bar{t}_i R_{dep}; \qquad \delta\bar{t}_i = \bar{t}_{i+1} - \bar{t}_i \\
&u_k = \sum_{i=1}^{n} u_{k,i} [H(\bar{t}_{i+1}) - H(\bar{t}_i)] \quad k \in \{1,2,3\} \\
&900 \leq T_s(u_k) \leq 1300 \qquad R_{dep} = k_a C_{A'} \;\; at \;\; \gamma
\end{aligned}
\tag{11}
$$

where F is the objective functional, $\mathcal{T}$ is the thickness of film at the end of the process, $\mathcal{T}_{obj}$ is the target thickness of the film, R_{dep} is the deposition rate of species a, $\mathcal{R}$ is the surface roughness of the deposited film, $\delta\bar{t}_i$ is the time-interval for i^{th} switching, T_s is the surface temperature and u_k is the magnitude of actuation. 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. 3, 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. 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 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 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.

5. RESULTS

The optimization problem of Eq.11 was solved in two steps. Initially, spatial uniformity of the deposited film was the only optimization objective (i.e. $w_1 \neq 0, w_2 = 0$). Subsequently, the microscopic objective was included into the optimization. The target film thickness, $\mathcal{T}_{obj}$ was $5{\times}10^{-6}$ m. Figure 2a 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 non-uniformity, defined as $\sqrt{\int (\mathcal{T} - \mathcal{T}_{obj})^2 / \mathcal{T}_{obj}}$, for ABB inlet configuration was found to be 87.66% and 57.72% for substrate temperature 1300K and 900K, respectively. The corresponding numbers were 215% and 188% respectively for BAB inlet configuration. For the optimal process operation, radial non-uniformity in the film was successfully reduced to 1.14% and 1.5% for the former and latter cases respectively. The corresponding inlet switching times where $757s$ and $703s$, and the substrate temperature profiles before and after switching are shown in Figure 3a 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 Figure 2b. It should be noted that BAB configuration with $T_s = 1300$ would result in the film with lowest surface roughness, however such operation is not optimal with respect to spatial film-thickness uniformity.

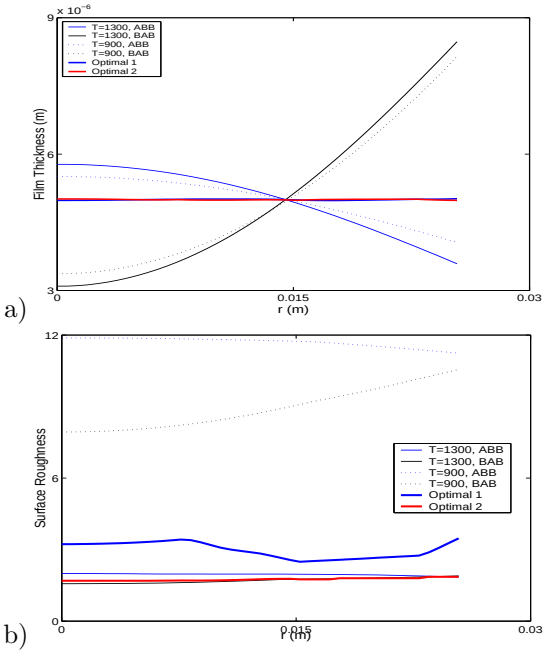

Fig. 2. Comparison of (a) deposition rate and (b) roughness profiles across the wafer surface, with macroscale (Optimal 1) and multiscale objective (Optimal 2).

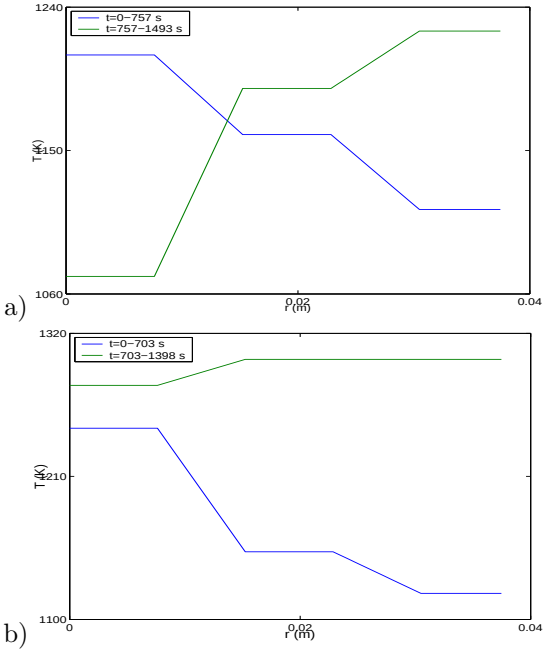

Fig. 3. Initial (blue line) and final (green line) optimal surface temperature profiles, for (a) macroscale only and (b) multiscale objective.

REFERENCES

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

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

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

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

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

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

Floudas, C. A. and P. M. Pardalos (1992). *Recent advances in global optimization.* Princeton University Press. Princeton, N.J.

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

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

Kevrekidis, I. G., C. W. Gear, J. M. Hyman, P. G. Kevrekidis, O. Runborg and K. Theodoropoulos (2003). Equation-free multiscale computation: enabling microscopic simulators to perform system-level tasks. *Comm. Math. Sciences* **1**, 715–762.

Li, J., D. Liao and S. Yip (1999). Nearly exact solution for coupled continuum/MD fluid simulation. *J. Comp. Mat. Des.* **6**, 95–102.

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

Meyer, C. A., C. A. Floudas and A. Neumaier (2002). Global optimization with nonfactorable constraints. *Ind. Eng. Chem. Res.* **41**, 6413–6424.

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

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

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